Idiosyncratic Information and Expected Rate of Returns

Schreder, Max Josef

Awarding institution:
King's College London

IDIOSYNCRATIC INFORMATION AND EXPECTED RATE OF RETURNS

DISSERTATION IN PART FULFILMENT OF THE REQUIREMENTS FOR THE ACADEMIC DEGREE DOCTOR OF PHILOSOPHY (PHD)

Max SCHREDER
King’s Business School
King’s College London
30 Aldwych, Bush House
WC2B 4BG London

Declaration

I hereby declare that the work presented in this thesis is original and entirely my own except where other sources are clearly cited. No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or institute.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement. I confirm that the following thesis does not exceed the word limit prescribed in the College regulations.

Name: Max Schreder
Date: 20th October 2017

Abstract

This thesis is situated at the interface between asset pricing and market microstructure theory. Motivated by seminal work of David Easley and Maureen O'Hara (2004), who present a cohesive framework in which idiosyncratic information impact firms' cost of equity, I contribute three interrelated research papers to this stream of research. My first paper "Idiosyncratic Information and Expected Rate of Returns: A Meta-Analytic Review of the Literature" provides a quantitative review of the literature examining the association between firm-specific information and expected rate of returns. Findings therein motivate my other two empirical papers.

My second paper "The Impact of Idiosyncratic Information on Expected Rate of Returns: A Structural Equation Modelling Approach" relates to work which tests the empirical validity of information-based return models and examines the question as to what extent a firm's information environment affects its cost of equity (CoE). Using a structural equation modelling approach—which is novel—it is shown that companies with high (low) quality information environments enjoy relatively lower (higher) CoE than otherwise identical firms; however, findings also show that the significance of this impact decreases with firm size, maturity and profitability as well as market competition.

My third paper "Implied Cost of Capital and Cross-Sectional Earnings Forecasting Models: Evidence from Newly Listed Firms" pertains to work on implied cost of capital (ICC)—which is part of a greater literature on expected rate of returns—and analyses the degree to which earnings forecasting models can be used to derive valid ICC estimates for newly listed firms. Results show that combining the earnings model of Hou et al. (2012, HVZ) with the earnings persistence (EP) model of Li and Mohanram (2014) into one forecasting solution generates less forecast bias, higher earnings response coefficients and more valid ICC estimates vis-à-vis the HVZ, EP and RI (residual income) models stand-alone. This suggests that for smaller and younger firms more complex forecasting solutions are required to ensure reliability of model-based ICC estimates.

The concluding chapter synthesizes the main findings of this thesis, indicates potential avenues for future research and discusses implications for practice.

Acknowledgements

Without the continuous help and support from many colleagues, friends and family, this thesis would have never been completed. I am particularly indebted to my PhD supervisors Colin Clubb and Tarik Driouchi for their invaluable guidance and encouragement throughout my studies. I am also deeply grateful to my parents—Max and Angelika—and my dear friends—Emre Abale, Aditi Gupta and Sebastian Kreis—who helped me back on my feet, whenever I stumbled.

Part of this thesis was written during my time visiting Cass Business School, where helpful discussions with Pawel Bilinski, Zhongwei Huang and Sonia Konstantinidi improved this thesis substantially. I also thank BAFA 2017 conference participants in Edinburgh for helpful comments, with special thanks to Mark Clatworthy and Elisabeth Dedman for constructive feedback. I thank Martin Edwards for his support on the SEM methodology, Hichem Khlif for helpful discussions on meta-analysis techniques, and Alexander Barinov for kindly providing FVIX data.

I also thank King's College London, School of Management and Business (now King's Business School) for awarding me with a full scholarship—this made my PhD studies so much easier. Financial support for part of this research was also gratefully received from the Faculty of Social Science and Public Policy Research Fund of King's College London.

Table of Contents

Figures, Tables and Appendices

Tables of Figures ……………………………………………………...…… page

Table of Tables ………………………………………………………………….. page

Table of Appendicies …..……………………………………………………….. page

1 Preface

The expected rate of return of a firm's equity (also referred to as cost of equity) is one of the most crucial numbers in corporate finance. Albeit its appropriate determination is an ongoing debate, its influence on corporations is clear: lower costs of equity, lead to higher valuations (i.e., higher stock prices) which is tantamount to increased shareholder wealth. Therefore, it is not surprising that academics, practitioners and policy maker alike aim to present ever new evidence as to how corporations can decrease their cost of equity (CoE).

Particularly active in this regard is a strand of research that combines insights from asset pricing theory (i.e., research on what return can be expected from an asset) and market microstructure theory (i.e., research on the drivers of price formation of an asset) to examine the impact of idiosyncratic information on firms' CoE. In general, a firm's information environment is broadly characterised by the following three attributes: (1) information quantity is the sheer amount of available public and private information about the future prospects of the firm; (2) information precision refers to the accuracy of available information about the future value of the company (e.g., within what margin of error can one infer future payoffs to shareholders from provided information) and (3) information asymmetry defines the degree to which investors are differentially informed (i.e., the fraction of privately informed vs. publicly informed investors). The interplay between these three attributes characterises the overall quality of a firm's information environment. Eventually, the extent to which these information attributes constitute an economically significant effect on the CoE is at the forefront of this line of inquiry.[1]

David Easley and Maureen O'Hara (2004, hereafter: EO) pioneer this realm in that they are the first to present a cohesive framework in which the effect of information quantity, precision and asymmetry on a firm's CoE is conceptualised. Their model demonstrates that CoE decreases with high-precision, low-asymmetry information environments and motivates a large body of theoretical and empirical work. Lambert, Leuz and Verrecchia (2012, hereafter: LLV) present an important extension of the EO model. They show

[1] From the outset it is important to note that some studies use the term *information quality* and *information precision* interchangeably; however, this study regards information precision to be only one of three attributes that impact the overall *quality of the information environment*. Appropriate terminology throughout this thesis makes this distinction clear.

that if markets are perfectly competitive (i.e., highly liquid), information asymmetry has no impact on the CoE over and above its impact on information precision; however, if markets are imperfectly competitive, information asymmetry might constitute a separate CoE effect. This subtle extension of the EO model entails that the degree of market competition represents a crucial conditioning variable in empirical settings. Taken together, both models make a strong stance as to why, how and when firm specific information should have a direct effect on CoE and, therefore, support the notion of information attributes being priced in the markets—despite their idiosyncratic nature.

That this notion is much debated (e.g., Hughes et al. (2007), Lambert et al. (2007)) is somewhat unsurprising, not least because any evidence supporting above models' predictions directly undermines neoclassic asset pricing theory which states that the only risk which impacts a firm's CoE is its exposure to (systematic) market risk, and any risk which is peculiar to the firm tends to be diversifiable; however, "if something is diversifiable is an empirical question that cannot be resolved by assumption within a theoretical model" (Shevlin, 2013, p. 448). Therefore, empiricism might be the final arbiter of this ongoing debate and further empirical work is warranted.

Any study examining the empirical veracity of the EO and LLV model must decide on how to proxy for both the cost of equity and the respective information attributes. Given on the one hand that the validity of the different expected return proxies is much debated, and on the other that proxies for information attributes are large in numbers as informed by both accounting and finance research, the empirical literature is voluminous and the conclusions reached vary widely depending on the proxies used by researchers. However, as highlighted in my first paper "Idiosyncratic Information and Expected Rate of Returns: A Meta-Analytic Review of the Literature" (Section 2) the following two observations are particularly startling.

First, extant work almost exclusively performs autonomous tests of the direct links between a single information attribute and CoE, which hinders the assessment of the relative importance of information precision, asymmetry and quantity as determinates of firms' expected returns (see paths *b*, *c* and *f* in Figure 1.1). Except for a recent study by

Bhattacharya et al. (2012)—who use path analysis to decompose the association between information precision and CoE into a direct and indirect path mediated by information asymmetry—such assessment is not provided in the literature. However, simultaneous analysis of the interplay between the attributes as well as an investigation of the direct, indirect and total effects of the attributes on firms' CoE is a necessary "first step" to disentangle the underlying complexity between idiosyncratic information and expected rate of returns (Beyer et al., 2010, p. 309).

Second, an aspiring literature—in particular in accounting research—uses the concept of implied cost of capital (ICC) to derive empirical measures of cost of equity.[2] A crucial input to the ICC methodology are reliable forecasts of future payoffs to shareholders, with analysts' consensus earnings forecasts commonly used in this regard. However, analysts' forecasts not only tend to be overly optimistic, which can result in severely upward biased ICC estimates, but are also mostly unavailable for young, small and financial distressed firms—the sort of firms which would be of "greatest interest to researchers examining issues related to information asymmetry, earnings quality, and disclosure where an ICC approach is used most often" (Li and Mohanram, 2014, p. 1153). To overcome these analyst associated deficiencies, Hou, van Dijk and Zhang (2012) and Li and Mohanram (2014) recommend the use of cross-sectional earnings forecasting models to predict future payoffs; nevertheless, the extent to which their models are applicable for the smallest, youngest and least followed firms in equity markets has not yet been examined.

My remaining two research papers address these limitations in extant work. The second paper "The Impact of Idiosyncratic Information on Expected Rate of Returns: A Structural Equation Modelling Approach" (Section 3) theoretically derives from market-microstructure research and relates to work which tests the empirical validity of information-based return models. The research question of this paper is as to what extent a

[2] The idea behind the ICC methodology is simple: use a specific valuation model, accept the current stock price as at least semi-strong efficient in the classical efficient market hypotheses sense (Fama, 1965, 1970) and *back-out* the internal rate of return which equates the current stock price of the firm with its expected future payoffs to shareholders (e.g. dividends). The internal rate of return is referred to as implied cost of capital (ICC) and regarded as the market participants' *ex ante* assessment of the firm's CoE.

firm's information environment affects its cost of equity (CoE). Using a structural equation modelling approach—which is novel—it is shown that companies with high (low) quality information environments enjoy relatively lower (higher) CoE than otherwise identical firms, with information precision and asymmetry being of equal CoE relevance, while information quantity effects being economically negligible. Figure 1.1 offers an illustration of the structural model examined as part of the paper's main analysis.

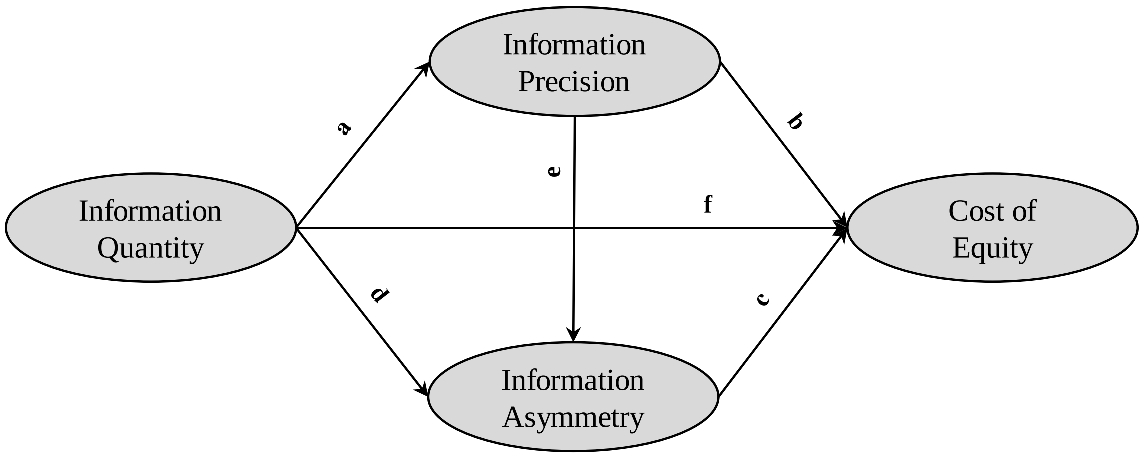


Figure 1.1: Structural Model of SEM

Oval figures indicate latent constructs which are only indirectly measurable through their impact on observable indicator variables (not shown in this figure). One-headed arrows between the constructs indicate regression relationships, and the end of each arrow denotes the dependent construct.

My third research paper "Implied Cost of Capital and Cross-Sectional Earnings Forecasting Models: Evidence from Newly Listed Firms" (Section 4) theoretically derives from asset pricing research and empirically pertains to work on ICC which is part of a greater literature on expected rate of returns. The research question in this paper examines degree to which earnings forecasting models can be used to derive valid ICC estimates for newly listed firms. Results show that combining the earnings model of Hou et al. (2012, HVZ) with the earnings persistence (EP) model of Li and Mohanram (2014) into one forecasting solution generates less forecast bias, higher earnings response coefficients and more valid ICC estimates vis-à-vis the HVZ, EP and RI (residual income) models stand-alone. This suggests that for smaller and younger firms more complex forecasting solutions are required to ensure reliability of model-based ICC estimates.

In the concluding chapter (Section 5), I synthesize the main findings of this thesis, indicate potential avenues for future research and discuss implications for practice.

2 Idiosyncratic Information and Expected Rate of Returns: A Meta-Analytic Review of the Literature

Abstract

This paper provides a quantitative review of the literature on the repercussions of idiosyncratic information on firms' expected rate of returns. In total, the results of 113 unique studies—which examine the cost of equity (CoE) effects of information *Precision*, *Asymmetry* and *Quantity*—are reconsidered. Results suggest that the association between firm-specific information and CoE is subject to moderate effects: first, a positive relationship between *Precision* and CoE is only significant in studies using *non*-accrual quality proxies for *Precision* and risk factor-based (RFB)/valuation model-based (VMB) proxies for CoE; second, almost all VMB studies confirm the positive association between *Asymmetry* and CoE, but there is notable variation in the conclusions reached when asset pricing tests are conducted; third, the link between *Quantity* and CoE is moderated by disclosure types and country-level factors in that firms' in comparatively weakly regulated countries tend to enjoy up to four times greater CoE benefits from more expansive disclosure than firms in strongly regulated markets.

2.1 Introduction

Extensive literature in accounting and finance investigates the extent to which idiosyncratic information affects price formation and return structures in capital markets. This strand of research commonly tests the proposition that firms with high (low) quality information environments should enjoy relatively low (high) expected rate of returns; more specifically, it is conjectured that firms can lower their expected returns if they disclose more value-relevant information to investors (*Quantity*), provide information of higher accuracy (*Precision*) and disseminate information more widely between investor groups (*Asymmetry*). While analytical work models these propositions elegantly (e.g., Diamond and Verrecchia (1991), Easley and O'Hara (2004), Lambert, Leuz and Verrecchia (2012)), the empirical validity of these models and their predictions is subject to debate (e.g., Core et al. (2008) Mohanram and Rajgopal (2009)).

Given on the one hand that the proper measurement of firms' expected rate of returns (*alias* cost of equity) is an ongoing debate in itself (e.g. Botosan and Plumlee (2005), Easton and Monahan (2016), Elton (1999)), and on the other hand that proxies for the information attributes are large in numbers as informed by both accounting and finance research, the empirical literature is voluminous and the conclusions reached vary widely depending on the proxies used by researchers. With that in mind, the main objective of this paper is to offer a systematic review of the extant literature to examine the reasons underlying the variation in results. More specifically, this review meta-analyses the links between *Precision*/*Quantity* and cost of equity (CoE), and provides a descriptive summary of extant findings on the association between *Asymmetry* and CoE.[3] To the best of my knowledge, this is the first study which quantitatively summarises the empirical literature on idiosyncratic information and expected rate of returns, and, as such, complements narrative literature reviews on this topic (e.g., Artiach and Clarkson (2011), Beyer et al. (2010), Healy and Palepu (2001), Kothari et al. (2016)).

[3] Given vast variation in research designs between studies examining the impact of *Asymmetry* on CoE (e.g., some studies use yearly, others monthly data; some focus on portfolio-level, other on firm-level) a meta-analysis is not feasible; hence, I focus on descriptive statistics only when examining this link.

Overall, findings suggest that the association between firm-specific information and CoE is subject to moderate effects as indicated by insignificant average effect sizes (*Precision*: -0.048 ± 0.095; *Quantity*: -0.066 ± 0.124) and a notable amount of studies (29%) which only find conditional results for the *Asymmetry* link.[4] Variation in the empirical measurement of both CoE and information attributes partially explains these mixed results: first, CoE effects of *Precision*, *Asymmetry* and *Quantity* are 3 to 6 times stronger in studies using risk factor-based (RFB)/valuation model-based (VMB) proxies than in studies conducing asset pricing tests or using realised returns as proxies for CoE; second, the controversy over the impact of *Precision* (*Asymmetry*) on firms' CoE stems—by and large—from the debate on the market pricing of accrual quality (PIN): whenever other proxies are used, results confirm the conjectured associations with CoE; similarly, *Quantity* results show that country-level factors and the type of information (financial vs. partial-/non-financial) mitigate disclosure effects on firms' CoE.

The rest of this paper proceeds as follows. Section 2.2 describes the conceptual framework in this study and derives the research hypotheses. Section 2.3 provides a succinct narrative review of previous empirical studies. Section 2.4 outlines the methodology. Section 2.5 reports results of the systematic literature review, and concluding remarks are provided in Section 2.6.

2.2 Conceptual Framework and Research Hypotheses

To facilitate the structure of this review, I present a conceptual framework based on which the studies in my sample are allocated (see Figure 2.1). The direct links between the information attributes and CoE are supported by analytical work. In a seminal paper, Easley and O'Hara (2004) demonstrate that a firm's CoE decreases with the accuracy of available information about the future value of the firm; that is, investors demand to be rewarded for bearing uncertainty about a firm's prospects which stems from imprecise information given to them, implying that firms which disclose higher quality information to investors can benefit from reduced CoE. Thus, I formulate Hypothesis 1:

[4] Average effect size plus/minus two standard deviations of the mean (i.e., 95% confidence interval).

H1: The higher (lower) the information precision of a firm, the lower (higher) its CoE.

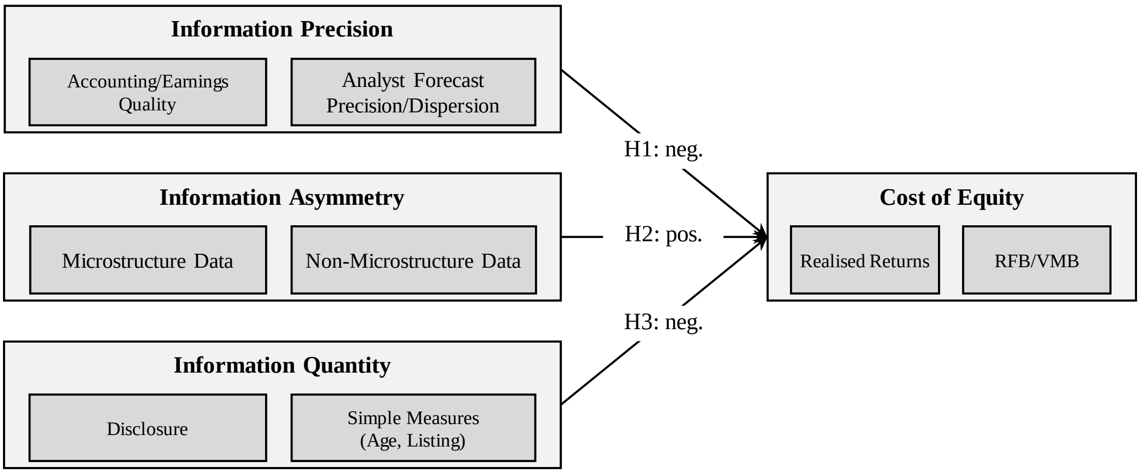


Figure 2.1: Conceptual Framework

This figure illustrates the direct links between the information attributes and CoE. For each attribute, the most commonly used empirical measures are indicated. VMB (RFB) Valuation model-based (Risk factor-based) cost of equity proxies.

Easley and O'Hara (2004) also show that as the fraction of uninformed investors as well as the number of private signals about the future value of the firm increases, its CoE decreases; that is, uninformed investors—who only have access to public information—require compensation for "losing out" against privately informed investors—who have access to both public and private information—when making investment decisions. The higher this informational disparity between these two groups is, the larger the CoE premium induced by *Asymmetry*. This corroborates Hypothesis 2:

H2: The higher (lower) the information asymmetry between investors, the higher (lower) the firm's CoE.

Finally, the estimation risk literature (e.g., Clarkson et al. (1996), Lewellen and Shanken (2002)) demonstrates that if the amount of information about a firm is low, investors have difficulties to accurately estimate the return and cash flow parameters of this particular firm (i.e., parameter uncertainty); this makes it a riskier investment vis-à-vis otherwise comparable firms, and, hence, induces higher CoE. These insights are summarised in Hypotheses 3:

H3: The larger (smaller) the quantity of available information about a firm, the lower (higher) its CoE.

The proceeding paragraphs provide a succinct narrative review of some notable studies that draw upon these hypotheses. This pinpoints prevailing debates in extant work, reveals commonly used proxies for *Precision*, *Asymmetry*, *Quantity* and CoE, and offers guidance towards the creation of meaningful sub-groups for later analyses.

2.3 Previous Empirical Studies

Firms' cost of equity is the dependent variable in the studies reviewed below and its empirical measurement varies widely between papers. While some researches rely on realised returns to proxy for CoE (e.g., Doukas et al. (2006) Konchitchki et al. (2016)), others use traditional risk factor-based estimates (e.g., Barth et al. (2013), Cohen (2008)), and yet others apply valuation model-based measures (e.g., Bhattacharya et al. (2012), Hou (2015)). Those studies which conduct asset pricing tests or use future realised returns as proxies of CoE are classified as realised return-based (REAL) in this study. The remaining "non-REAL" papers are further divided into two categories: studies using risk factor-based (RFB) estimates—such as rCAPM, rFF4 (Carhart (1997), Fama and French (1993, 2015), Lintner (1965), Mossin (1966), Sharpe (1964))—and papers applying valuation model-based (VMB) proxies, i.e., implied cost of capital (ICC) estimates (see Botosan and Plumlee (2002), Claus and Thomas (2001), Easton (2004), Gebhardt et al. (2001), Gode and Mohanram (2003), Gordon and Gordon (1997), Ohlson and Juettner-Nauroth (2005) for commonly used ICC estimates).[5]

Empirical measures are inherently flawed (Rao, 1973), and so are the CoE proxies mentioned above. For instance, Elton (1999) in his presidential address states: "The use

[5] The intuition behind the ICC framework is straightforward: use a specific valuation model, accept the current stock price as at least semi-strong efficient in the classical efficient market hypotheses sense (Fama 1965, 1970) and "back-out" the internal rate of return which equates current stock price of the firm with its expected future payoffs to shareholders, where future payoffs are commonly proxied by analysts' earnings forecasts. The internal rate of return is then considered the market participants' *ex ante* assessment of the firm's CoE.

of average realized returns as a proxy for expected returns relies on a belief that information surprises tend to cancel out over the period of a study and realized returns are therefore an unbiased estimate of expected returns. However, I believe that there is ample evidence that this belief is misplaced" (p. 1199). In a similar vein, Fama and French (1997) argue that RFB proxies are "woefully imprecise estimates of the cost of equity" (p. 154), given that these estimates are based on noisy past realised returns; that is, the unexpected news component in realised returns tend to corrupt the reliability of factor loading and factor premia estimates in RFB models. Finally, Easton and Monahan (2005) examine a number of different ICC estimates and conclude that—due to the optimism in analysts' earnings forecasts—"these proxies are unreliable" (p. 501).

Despite these concerns, two major conclusions regarding the empirical validity of the different measures seem to be justified. First, proxying for CoE by realised returns is problematic in that discount rate and cash-flow news have a significant impact on firm-level returns: realised returns are noisy estimates of CoE (e.g., Chen et al. (2013), Ogneva (2012), Vuolteenaho (2002)). Second, VMB proxies show somewhat greater construct validity than RFB proxies in terms of association with future realised returns, common risk-factors, and predictive power of returns (Botosan and Plumlee (2005), Botosan et al. (2011), Lee et al. (2010, 2015)). Taken together, this suggests that researchers' choice of CoE measurement might distort findings and partially explain mixed results across studies when analysing the impact of *Precision*, *Asymmetry* and *Quantity* on firms' expected returns.

2.3.1 Information Precision and Cost of Equity

The literature analysing the hypothesis of lower information accuracy resulting in higher CoE (H1) is thoroughly researched (e.g., 38 papers are assigned to this link in my analysis) and can be categorised in two major strands: *accounting/earnings quality* and s*ecurity analyst forecast*-based studies.

The first stream is pioneered by Francis, La Fond, Olsson and Shipper (2004, 2005). The authors examine the association between the quality of accounting information and investors' return expectations, and demonstrate that firms' CoE decrease as their earnings quality measures increase (e.g., accrual quality, value relevance); however, in an influential paper, Core, Guay and Verdi (2008) strongly question the validity of this early results. Despite continuous evidence for the proposition that earnings/accounting quality is inversely related to expected returns (e.g., Aboody et al. (2005), Barth et al. (2013), Hou (2015)), this link remains challenged in a number of papers (Cohen (2008), Khan (2008), (McInnis, 2010)).

The second stream of research uses security analyst forecasts to proxy for information precision; it is argued that the less uncertainty exists about the prospects of a firm, the greater the consensus among security analysts, which is then reflected in more precise forecasts (Barry and Brown (1985), Barron and Stuerke (1998), Barron et al. (1998)). Botosan and Plumlee (2013, working paper in 2003) provide initial support that total analyst forecast precision (AFP) is negatively associated with CoE. Botosan et al. (2004) decompose total analyst forecast precision into its private and public information components and show that public (private) AFP is negatively (positively) associated with CoE. Interestingly, Barron et al. (2012) show that for firms with limited public information, private analyst forecast precision is in fact *negatively* associated with CoE, which might be explained by the dual effect of private information precision. On the one hand, more precise private information increases information asymmetry which induces higher CoE (e.g., Diamond and Verrecchia (1991)), but on the other hand it might also decrease CoE when the quantity of public information is limited; that is, private information is still better than no information at all (Easley and O'Hara (2004), Lambert et al. (2012)).

In summary, accounting quality—and more specifically earnings quality—appears to be an important determinant of firms’ information precision; however, since the study of Core et al. (2008) its association with CoE is subject to controversy. What is more, CoE seems to be decreasing with total analyst forecast precision, but its relation is less clear regarding the private and public components of it.

2.3.2 Information Asymmetry and Cost of Equity

Numerous papers investigate the proposition that CoE increases with information asymmetry (H2); that is, the greater the informational disadvantages between privately informed and publicly uninformed investors, the higher firms' CoE. Twenty-two (22) papers are included in my analysis, with *Asymmetry* proxies mainly deriving from market microstructure data (e.g., PIN scores; bid-ask spreads).[6] Easley, Hvidkjaer and O'Hara (2002) are the first to document a positive association between CoE and PIN scores. Mohanram and Rajgopal (2009) question these findings and conclude that "there is not much evidence to support the interpretation that information risk, proxied by PIN, is a source of priced information risk (p. 241)". However, the authors acknowledge that while their paper suggests "PIN is not priced risk, it is difficult to make more general statements about the pricing of information risk since information risk can […] be proxied by different empirical variables." In fact, using spread-based proxies, Amihud and Mendelson (1986), Bhattacharya et al. (2012), Levi and Zhang (2015)—among others—document a significantly positive association with expected returns.

Taken together, market microstructure-based estimates are common proxies for *Asymmetry* in empirical work. While there is an ongoing debate about the pricing of PIN (e.g., Botosan and Plumlee (2013), Duarte and Young (2009)), the inverse relationship between bid-ask-spreads and firms' CoE seems to be widely accepted.

2.3.3 Information Quantity and Cost of Equity

A large literature examines the association between the quantity of available information about a firm and its CoE (see H3), with 56 papers being analysed in this study. One stream of research operates simple proxies—such as firm age or period of listing—as measures of information quantity and demonstrates that these proxies are negatively correlated with CoE (Barry and Brown (1984, 1985), Clarkson and Thompson (1990), Kumar et al. (2008), Zhang (2006)).

[6] PIN scores measure the probability that the next trade order is from a privately informed investor; i.e., based on private information (Brown et al., 2004). Bid-ask spreads are a measure of the adverse selection problem market makers are exposed to and assumed to increase with information asymmetry (Copeland and Galai (1983), Glosten and Milgrom (1985)).

A second stream of research uses firms' disclosure levels as a proxy for *Quantity*, with ample evidence supporting the predicted negative association with CoE (e.g., Baginski and Rakow (2012), Botosan (1997), Campbell et al. (2014), Cao et al. (2017), Healy et al. (1999), Francis et al. (2005b), Fu et al. (2012), Kothari et al. (2009), Ng and Rezaee (2015)). However, a large variety of different research designs, disclosure types examined and disclosure metrics used by researchers' makes it difficult to generalise results in narrative reviews; that is, some studies investigate mandatory disclosure (e.g., Campbell et al. (2014), Core et al. (2015)), while others concentrate on voluntary disclosure aspects (e.g., Botosan and Plumlee (2002), Francis et al. (2008)); some papers focus on financial disclosure (e.g., Baginski and Rakow (2012), Evans (2016)), while others examine non-financial disclosure (e.g., Dhaliwal et al. (2011), Ng and Rezaee (2012)); and, some authors use self-constructed disclosure scores (e.g., Botosan (1997), Kothari et al. (2009)), while others rely on commercially available ones (e.g., Healy et al. (1999), Richardson and Welker (2001)), and yet others use simple dummy variables to distinguish between disclosing and non-disclosing firms (e.g., Ogneva et al. (2007), Cao et al. (2017)).

Overall, the disclosure literature dominates "simple-proxy" studies in terms of volume and recognition whenever the link between *Quantity* and CoE is tested. However, it should be noted that commonly used proxies for disclosure levels—be they self-constructed or externally provided—tend to measure disclosure along both a quantity and quality dimension, which makes them noisy measures of *Quantity*.[7] What is more, whilst there is strong evidence supporting an inverse relation between disclosure and firms' CoE in general, more evidence—along the lines of Richardson and Welker (2001) and Mangena et al. (2016)—on the types of disclosure which are particularly CoE relevant seems beneficial. This study contributes such evidence.

2.4 Methodology

The main objective of this paper is to provide a meta-analytic review of the literature on idiosyncratic information and expected rate of returns. However, only for H1 and H3

[7] For instance, Cheng et al. (2006, p. 179) state that "while prior empirical research has used the *quantity* of disclosure as a proxy for the *quality* of disclosure quality, in many cases disclosure quantity and quality are not separable information attributes."

research designs between studies are homogenous enough to carry out such analyses; studies relating to H2 are often lacking necessary information (i.e., sample size is inconsistently reported) and data comparability (e.g., some studies use yearly, others monthly data; some focus on portfolio-level, other on firm-level) to conduct a meaningful meta-analysis; thus, for H2 I only provide descriptive statistics (e.g., fraction of studies confirming/rejecting the hypothesis; proportion of what proxies are used).

2.4.1 Data Collection

In line with research hypotheses H1-H3, I search for different combinations of several keywords (see Table 2.1) in the following databases to identify relevant studies: ISI Web of Science; ScienceDirect; Emerald Management Ejournals; SSRN. A review of all top-tier journals in accounting and finance as well as the reference list of all identified articles complements my search.[8] Analytical papers (e.g., Bertomeu et al. (2011), Cheynel (2013), Christensen et al. (2010), Dutta and Nezlobin (2017), Hughes et al. (2007), Lambert et al. (2007), Strobl (2013)) and studies lacking statistical information required for the meta-analysis of H1 and H3 (e.g., Aboody et al. (2005), Beneish et al. (2008), Chen et al. (2007), Clement et al. (2003), Hirshleifer et al. (2012)) are excluded during this process.

The literature applies various different measures for *Precision*, *Asymmetry, Quantity* and CoE. Based on the most commonly used proxies, I create several sub-groups for each information attribute and CoE (see Figure 2.1). Each study in my sample is coded accordingly: for instance, Francis et al. (2004) examine the impact of seven different earnings quality measures on two implied cost of capital proxies; hence, I categories this study as using "Accounting Quality" proxies for information precision and "VMB" proxies for CoE. A number of insightful papers are excluded from the analysis, since the novelty of the proxies used in these studies hinders their allocation to one of my sub-groups: for example, El Ghoul et al. (2013) put forward firms' geographic distance from financial

8 *Accounting Journals:* Accounting Review; Accounting, Organizations and Society; Journal of Accounting and Economics; Journal of Accounting Research; Contemporary Accounting Research; Review of Accounting Studies; *Finance Journals:* Journal of Finance; Journal of Financial Economics; Review of Financial Studies; Journal of Corporate Finance; Journal of Financial and Quantitative Analysis; Journal of Financial Intermediation; Journal of Money, Credit and Banking; Review of Finance.

centres as a measure of information asymmetry and show that CoE decreases with proximity; Cao et al. (2015) find that company reputation is inversely related to CoE, and Muino and Trombetta (2009) document a significant impact of distorted graph disclosure on expected returns. Furthermore, the focus of this review lies on firm-level results, therefore, I exclude country-level studies from the sample (e.g., Bhattacharya et al. (2003), Li (2015)).

Table 2.1: Attributes and Keywords

Attributes	**Keywords**
Information Precision	▪ information precision/accuracy/quality ▪ financial reporting quality ▪ accounting quality ▪ earnings quality/management ▪ earnings attributes ▪ accrual quality ▪ analyst forecast ▪ security analyst ▪ analyst forecast precision/accuracy/dispersion ▪ earnings/analyst characteristic ▪ earnings persistence/predictability/smoothness/value relevance/timeliness/conservatism
Information Asymmetry	▪ information asymmetry/dissemination/dispersion ▪ informational dis/advantages ▪ un/informed investor ▪ public/private information ▪ probability of informed trading/PIN ▪ bid-ask spread ▪ investor concentration/competition ▪ market liquidity ▪ firm-specific information ▪ idiosyncratic information
Information Quantity	▪ information quantity/amount ▪ information/estimation risk ▪ mandatory/voluntary disclosure ▪ financial/non-financial disclosure ▪ age/listing/operating history ▪ media/press coverage ▪ firm/company prominence

Table continued next page.

Table 2.1: Attributes and Keywords (cont.)

Attributes	Keywords
Cost of Equity	▪ cost of equity ▪ cost of capital ▪ implied cost of capital ▪ expected rate of return ▪ required rate of return ▪ discount rate ▪ weighted cost of capital

The table shows for each information attribute and cost of equity the respective keywords searched for in several databases and journals.

2.4.2 Meta-Analysis Techniques

2.4.2.1 Effect Size

Meta-analysis techniques require the use of effect size which is the study's Pearson r correlation coefficient between the dependent and independent variable of interest (Hunter et al., 1982).[9] If a study only reports regression results, t-statistics are converted into r coefficients as $\sqrt{t^2/t^2 + df}$, where df is degrees of freedom. If only p-values are provided, corresponding Z scores can be obtained from the normal table, which are then transferred into r coefficients by using $Z/\sqrt{N}$, where N is sample size (Rosenthal, 1991).

Whenever a study uses multiple, but similar measures for a variable (say, different proxies for earnings quality), the study's average effect size is recorded; thus, these studies only appear once in the analysis (Ahmed and Courtis, 1999); however, if a study tests different proxies for different samples (say, one RFB and one VMB CoE sample), the effect size for each sub-sample is recorded separately; thus, these studies appear twice in the analysis. For example, 35 studies for the link between *Precision* and CoE contribute 48 observations to my analysis (Asymmetry: 22 studies and 28 observations; Quantity: 56 studies and 62 observations). Also, where necessary, I multiply a study's effect size by negative one to ensure consistent interpretation across proxies and to conform to the

9 I use Spearman correlations, if Pearson correlations are not reported.

intuition of the underlying hypotheses of this paper. This is necessary because interpretations of results vary depending on which proxies are used: some proxies are constructed such that *lower* values signal higher information quality, while others are constructed such that *higher* values also signal higher information quality. For instance, in most studies, lower earnings quality proxies indicate higher information precision—confirming H1 if a positive, not negative, correlation with CoE is observed—while higher analyst-based measures signal higher *Precision* and confirm H1 if a negative correlation with CoE is observed. In such instances, I multiply effect sizes (of earnings quality) by negative one to reverse and re-align the interpretation of results with the intuition of the underlying hypotheses (H1).

After the effect size for each study is calculated, the mean correlation ($\bar{r}$) for the population is estimated as shown in Eq. (2.1), where N_i is the sample size and r_i is the Pearson correlation coefficient for study *i*.

$$\bar{r} = \frac{\sum(r_i N_i)}{\sum N_i} \tag{2.1}$$

The population variance (S_p^2) is estimated as the difference between observed variance (S_r^2) and the sampling error variance (S_e^2) as shown in equations (2.2) to (2.4), where *K* is the total number of studies included in the analysis.

$$S_p^2 = S_r^2 - S_e^2 \tag{2.2}$$

$$S_r^2 = \frac{\sum N_i (r_i - \bar{r})^2}{\sum N_i} \tag{2.3}$$

$$S_e^2 = \frac{(1 - \bar{r}^2)^2 K}{\sum N_i} \tag{2.4}$$

Eventually, the 95% confidence interval is calculated as follows:

$$\left[\bar{r} - S_p Z_{0.975}; \bar{r} + S_p Z_{0.975}\right] = \left[\bar{r} - S_p(1.96); \bar{r} + S_p(1.96)\right] \tag{2.5}$$

2.4.2.2 Homogeneity Tests

To test for homogeneity in the data—that is, to examine if variation in results is due to sampling errors or moderating effects—two methods are followed. First, putting the sampling error variance into perspective to observed variance (S_e^2/S_r^2) reveals the degree to which the residual variance is trivial; that is, if more than 75 percent of the variation in results can be attributed to sampling error (the suggested cut-off in the literature), then the relation under investigation is considered to be homogenous and unmoderated (Ahmed and Courtis (1999), Khlif and Chalmers (2015), Pearlman et al. (1980)). Second, the chi-square test statistic shown in equation (2.6) is calculated, where high statistical significance rejects the null of homogeneity, indicating that moderating effects might impact results across studies (*K-1* is the degrees of freedom and N = $\sum N_i$). While statistically powerful, it should be noted that chi-square statistics are directly proportional to sample size (Fan et al., 1999), which makes it difficult to accept homogeneity at conventional levels in large sample studies like this one.

$$\chi_{K-1}^2 = \frac{NS_r^2}{(1 - \bar{r}^2)^2} \tag{2.6}$$

2.4.2.3 Sample Size

Most of the studies analysed report results for a multi-period sample. This poses the question of how to define sample size (i.e. N_i) in equations (2.1) to (2.6). Using the number of unique firms appears to be most appealing; however, this information is often missing. Conversely, the number of firm-years (firm-quarters or firm-months) is reported most consistently, but might bias meta-analytic results towards studies spanning longer sampling periods, without necessary covering more firms. I address this point as follows: when transforming t-statistics (p-values) from regression results into r coefficients, I determine the degrees of freedom (sample size) based on the number of observations in the regression (e.g., firm-years). When summarising among studies (e.g., computing mean effect size, population variance, chi-square statistic) the average number of firms (e.g., firm-years divided by number of years) is used as sample size. This maintains the internal integrity of each study—larger sample studies generate robuster results—while ensuring a “sampling-period-independent” impact on the meta-level.

2.5 Meta-Analytic Results

Thirty-five (35) studies in my sample examine the link between *Precision* and CoE, 22 the link with *Asymmetry* and 56 the link with *Quantity*. Given that some papers provide findings for multiple sub-categories (for instance, RFB and VMB proxies), subsequent analyses are based on 48 observations for *Precision*, 28 for *Asymmetry* and 62 for *Quantity*. Hereafter I use "observations" and "studies/papers" interchangeably, but results always refer to the number of observations. The *Precision*, *Asymmetry* and *Quantity* studies included in my sample are shown in Table 2.2, Table 2.3, Table 2.4, respectively; Table 2.5 summarises this information and reports descriptive statistics across the entire sample.

Table 2.2: Information Precision on CoE – Description of Studies Included in the Meta-Analysis

Study	Journal[*]	Country[†]	No. of firms[§]	No. of firm-years[‡]	Sampling Period	DV: CoE[¥]	IV: Precision[£]	Direct Link?	Effect size (Pearson's *r* coefficient)	Source of Information
Artiach and Clarkson (2014)	AJM	US	196	3,138	1985-2000	VMB	Acc.Qual.	Confirm	-0.0593	Table 3, p. 305
Barron et al. (2012)	SSRN	US	307	8,606[a]	1983-2010	VMB	Analyst	Confirm	-0.2327	Table 2, p. 29
Barth et al. (2013)	JAE	US	1,985	51,612	1974–2000	RFB	Acc.Qual.	Confirm	-0.1210	Table 1, p. 214
Barth et al. (2013)	JAE	US	1,985	51,612	1974–2000	REAL	Acc.Qual.	Confirm	-0.0380	Table 1, p. 214
Berger et al. (2012)	SSRN	US	1,665	41,615	1980-2004	RFB	Acc.Qual.	Confirm	-0.0475	Table 7, p. 33
Berger et al. (2012)	SSRN	US	1,015	25,365	1980-2004	VMB	Acc.Qual.	Confirm	-0.0830	Table 5, p. 29
Bhattacharya et al. (2012)	TAR	US	1,054	12,648	1993-2005	RFB	Acc.Qual.	Confirm	-0.1770	In-text, p. 475
Bhattacharya et al. (2012)	TAR	US	1,054	12,648	1993-2005	VMB	Acc.Qual.	Confirm	-0.2243	Table 2, p. 463
Botosan and Plumlee (2013)	JBFA	US	555	6,656	1993-2004	VMB	Analyst	Mixed	-0.1470	Table 3, p. 1061
Botosan et al. (2004)	RAST	US	312	2,804	1993-2001	VMB	Analyst	Mixed	-0.0930	Table 2, p. 247
Callen et al. (2013)	CAR	US	1,129	29,345	1981-2006	REAL	Acc.Qual.	Confirm	-0.0354	Table 5 & 6, pp. 283-85
Callen et al. (2013)	CAR	US	841	19,336	1984-2006	VMB	Acc.Qual.	Confirm	-0.0416	Table 7, p. 287
Chan et al. (2009)	MF	UK	416	5,403	1987-1999	VMB	Acc.Qual.	Mixed	-0.0190	Table 3, p. 336
Chan et al. (2009)	MF	UK	416	5,403	1987-1999	REAL	Acc.Qual.	Mixed	-0.0288	Table 9, p. 342
Chan et al. (2016)	NAJEF	US	1,828	32,910	1996-2013	RFB	Analyst	Confirm	-0.1203	Table 4, p.125
Chen et al. (2008)	JAAF	US	2,122	53,048	1980-2004	REAL	Acc.Qual.	Confirm	-0.0420	Table 2, p. 480
Chen et al. (2008)	JAAF	US	614	15,339	1980-2004	VMB	Acc.Qual.	Confirm	-0.0433	Table 4, p. 489
Cohen (2003)	SSRN	US	1,111	16,664	1987-2001	VMB	Acc.Qual.	Reject	-0.0045	Table 6 & 8, pp. 44-6
Cohen (2008)	APJAE	US	1,074	18,264	1987−2003	VMB	Acc.Qual.	Reject	-0.0068	Table 3, p. 83
Cohen (2008)	APJAE	US	1,074	18,264	1987−2003	RFB	Acc.Qual.	Reject	-0.0052	Table 4, p. 85
Core et al. (2008)	JAE	US	2,909	93,093	1970-2001	REAL	Acc.Qual.	Reject	0.0003	Table 4 & 5, pp. 11-3
Core et al. (2008)	JAE	US	814	21,979	1975-2001	VMB	Acc.Qual.	Reject	-0.0386	Table 8, p. 18
Diether et al. (2002)	JF	US	2,908	66,884	1980-2002	REAL	Analyst	Reject	0.0089	Table 9, p. 2136
Diether et al. (2002)	JFQA	US	1,203	22,854	1983-2001	REAL	Analyst	Confirm	-0.0585	Table 8, p. 597
Eliwa et al. (2016)	IRFA	UK	587	4,112	2005–2011	VMB	Acc.Qual.	Confirm	-0.0714	Table 1, p. 131

Table continued next page.

Table 2.2: Information Precision on CoE – Description of Studies Included in the Meta-Analysis (cont.)

Study	Journal*	Country†	No. of firms§	No. of firm-years‡	Sampling Period	DV: CoE¥	IV: Precision£	Direct Link?	Effect size (Pearson's *r* coefficient)	Source of Information
Francis et al. (2004)	TAR	US	790	21,334	1975-2001	VMB	Acc.Qual.	Confirm	-0.0481	Table 5, 6 & 9, pp. 990-1001
Francis et al. (2005a)	JAE	US	1,722	55,092	1970-2001	VMB	Acc.Qual.	Confirm	-0.0248	Table 2, p. 309
Garcia Lara et al. (2011)	RAST	US	1,875	54,389	1975-2003	REAL	Acc.Qual.	Confirm	-0.0139	Table 4 & 6, pp. 261-4
Gray et al. (2009)	JBFA	AU	170	1,362	1998-2005	VMB	Acc.Qual.	Confirm	-0.0676	Table 3, p. 63
Gray et al. (2009)	JBFA	AU	170	1,362	1998-2005	REAL	Acc.Qual.	Confirm	-0.0730	Table 4 & 5, pp. 66-8
Hou (2015)	RAST	US	1,418	41,134	1982-2010	VMB	Acc.Qual.	Confirm	-0.1275	Table 2, p. 1073
Hwang and Lim (2012)	APJFS	US	645	9,672	1993-2007	VMB	Acc.Qual.	Confirm	-0.1640	Table 2, p. 471
Kim and Qi (2010)	TAR	US	2,802	103,682	1970-2006	REAL	Acc.Qual.	Mixed	-0.0040	Table 4, pp. 947-9
Kim and Sohn (2013)	JAPP	US	1,211	30,276	1987-2011	VMB	Acc.Qual.	Confirm	-0.0673	Table 2, p. 529
Konchitchki et al. (2016)	RAST	US	2,567	100,095	1976-2014	REAL	Acc.Qual.	Confirm	-0.0118	Table 5, p. 21
Larson and Resutek (2015)	SSRN	US	79	2,684	1978-2011	REAL	Acc.Qual.	Mixed	-0.0491	Table 6 & 8, p. 42-4
Larson and Resutek (2015)	SSRN	US	49	1,728	1977-2011	VMB	Acc.Qual.	Mixed	-0.1102	Table 9, p. 45
Latiff and Taib (2011)	ATBMR	MY	141	141	2004	VMB	Acc.Qual.	Confirm	-0.1624	Table 4, p. 6
Liu and Wysocki (2016)	QJF	US	1,454	68,348	1960-2006	RFB	Acc.Qual.	Mixed	-0.1200	Table 3, p. 15
Liu and Wysocki (2016)	QJF	US	945	44,392	1960-2006	VMB	Acc.Qual.	Mixed	-0.0700	Table 3, p. 15
Mashruwala and Mashruwala (2011)	TAR	US	2,561	92,187	1971-2006	REAL	Acc.Qual.	Mixed	-0.0048	Table 6, p. 1368
McInnis (2010)	TAR	US	1,777	56,870[b]	1975-2006	REAL	Acc.Qual.	Reject	-0.0025	Table 1, p. 321
McInnis (2010)	TAR	US	438	14,008	1975-2001	VMB	Acc.Qual.	Reject	-0.0444	Table 4, p. 328
Ogneva (2012)	TAR	US	2,184	80,790	1970-2006	REAL	Acc.Qual.	Mixed	-0.0048	Table 4, pp. 1433-4
Othman (2012)	AJBA	MY	461	3,688	2000-2007	VMB	Acc.Qual.	Confirm	-0.0319	Table 3, p. 17
Safdar and Yan (2016)	CFRI	CN	1,251	8,754	2006-2012	VMB	Acc.Qual.	Mixed	-0.0244	Table 2, p. 87
Safdar and Yan (2016)	CFRI	CN	1,251	8,754	2006-2012	REAL	Acc.Qual.	Mixed	-0.0138	Table 3, 5, & 6, p. 89-92
Sheng and Thevenot (2015)	IJF	US	128	3,583[a]	1984-2011	VMB	Analyst	Confirm	-0.3300	Table 2, p. 521

Notes: * *Journal names along with their ABS 2015 ranking are shown in Appendix 2.1.* † *AU: Australia; CN: China; MY: Malaysia; UK: United Kingdom; US: United States.* § *Number of firms is approximated as number of firm-years divided by number of sample years.* ‡ *When multiple samples are selected for one study, average sample size is reported.* ¥ *REAL: realised-return, RFB: risk factor-based, VMB: valuation model-based cost of equity proxy.* £ *Acc.Qual: Accounting/Earnings Quality, Analyst: security analyst forecast based proxy.* **a** *Converted firm-quarters into firm-years.* **b** *Converted firm-months into firm-years.*

Table 2.3: Information Asymmetry on CoE – Description of Studies Included in the Analysis

Study	Journal*	Country†	Sampling Period	DV: CoE¥	IV: Asymmetry£	Proxy‡	Direct Link?	Source of Information§
Akins et al. (2012)	TAR	US	1984-2009	REAL	Micro.	B/A	Mixed	Table 3, p. 48
Akins et al. (2012)	TAR	US	1984-2005	REAL	Micro.	PIN	Mixed	Table 4, p. 50
Amihud and Mendelson (1986)	JFE	US	1961-1980	REAL	Micro.	B/A	Confirm	Table 3, p. 236
Armstrong et al. (2011)	JAR	US	1988-2006	REAL	Micro.	B/A	Mixed	Table 3 & 4, pp. 23-8
Armstrong et al. (2011)	JAR	US	1976-2006	REAL	Non-Micro.	Acc.Qual.	Mixed	Table 3 & 4, pp. 23-8
Armstrong et al. (2011)	JAR	US	1976-2006	REAL	Non-Micro.	Analyst	Mixed	Table 3 & 4, pp. 23-8
Aslan et al. (2011)	JEF	US	1965-2009	REAL	Micro.	PIN	Confirm	Table 8, p. 796
Barron et al. (2012)	SSRN	US	1983-2010	VMB	Non-Micro.	Analyst	Confirm	Table 4, p. 32-4
Bhattacharya et al. (2012)	TAR	US	1993-2005	VMB	Micro.	B/A	Confirm	Table 2, p. 463
Bhattacharya et al. (2012)	TAR	US	1993-2005	VMB	Micro.	PIN	Confirm	Table 2, p. 463
Botosan and Plumlee (2013)	JBFA	US	1993-2004	VMB	Micro.	PIN	Confirm	Table 4, p. 1062
Botosan and Plumlee (2013)	JBFA	US	1993-2004	REAL	Micro.	PIN	Reject	Table 4, p. 1062
Brennan et al. (2016)	MS	US	1983-2010	REAL	Micro.	PIN	Confirm	Table 4 & 5, pp. 2469-70
Choi et al. (2016)	SSRN	CN	1996-2007	REAL	Non-Micro.	Comp.	Confirm	Table 6, p. 36
Duarte and Young (2009)	JFE	US	1984-2004	REAL	Micro.	PIN	Reject	Table 10, p. 136
Duarte et al. (2008)	JFE	US	1985-2000	REAL	Micro.	PIN	Confirm	Table 5, p. 37
Easley et al. (2002)	JF	US	1984-1998	REAL	Micro.	PIN	Confirm	Table 6, p. 2213
Easley et al. (2010)	JFQA	US	1983-2001	REAL	Micro.	PIN	Confirm	Table 6, p. 307
Eleswarapu and Reinganum (1993)	JFE	US	1961-1990	REAL	Micro.	B/A	Mixed	Table 2, p. 379
He et al. (2013)	IREF	AU	2001-2008	VMB	Micro.	B/A	Confirm	Table 4 & 5, p. 617-8
Hwang et al. (2013)	JAE	KR	2000-2004	VMB	Micro.	PIN	Mixed	Table 5, p. 158
Hwang et al. (2013)	JAE	KR	1995-2005	REAL	Micro.	PIN	Mixed	Table 9, p. 162
Kang (2010)	JBF	US	1984-2002	REAL	Micro.	PIN	Mixed	Table 4, p. 2990
Levi and Zhang (2015)	MS	US	1993-2003	REAL	Micro.	B/A	Confirm	Table 3, p. 361

Table continued next page.

Table 2.3: Information Asymmetry on CoE – Description of Studies Included in the Analysis (cont.)

Study	Journal*	Country†	Sampling Period	DV: CoE¥	IV: Asymmetry£	Proxy‡	Direct Link?	Source of Information§
Luong et al. (2011)	SSRN	US	1984-2006	REAL	Micro.	PIN	Mixed	Table 5, pp. 46-7
Mohanram and Rajgopal (2009)	JAE	US	1984-2002	VMB	Micro.	PIN	Reject	Table 7, p. 239
Mohanram and Rajgopal (2009)	JAE	US	1984-2002	REAL	Micro.	PIN	Reject	Table 1, p. 230
Yan and Zhang (2014)	JBF	US	1983-2005	REAL	Micro.	PIN	Confirm	Table 8, p. 147

Notes: * *Journal names along with their ABS 2015 ranking are shown in Appendix 2.1.* † *AU: Australia; CN: China; KR: South Korea; US: United States.* ¥ *REAL: realised-return, VMB: valuation model-based cost of equity proxy.* £*(Non-)Micro.: (Non-)Microstructure-based proxies.* ‡ *B/A: Bid/Ask-spreads; PIN: Probability of informed trading scores; Acc.Qual.: Accounting/Earnings Quality, Comp.: market competition, Analyst: security analyst forecast based proxy.* § *Information used to decide if a study confirms/rejects or finds mixed results regarding the direct link between Asymmetry and CoE.*

Table 2.4: Information Quantity on CoE – Description of Studies Included in the Meta-Analysis

Study	Journal[*]	Country[†]	US?[†]	Disc. Env.[€]	Disc. Reg. Score[$]	No. of firms[§]	No. of firm-years[‡]	Sampling Period	DV: CoE[¥]	IV: Quantity[£]	Disc. Type[£]	Disc. Metric[£]	Direct Link?	Effect size (Pearson's *r* coefficient)	Source of Information
Al Guindy (2016)	SSRN	US	Y	HDE	1.00	1,232	8,626	2007-2013	VMB	Disc.	NF	Dummy	Confirm	-0.0870	Table 5, p. 30
Al Guindy (2016)	SSRN	US	Y	HDE	1.00	54	381	2007-2013	VMB	Disc.	FF	Dummy	Confirm	-0.1008	Table 6, p. 31
Al-Hadi et al. (2015)	JMFM	6 Count.	N	LDE	n/a	141	705	2007-2011	VMB	Disc.	PF	SCI	Confirm	-0.0160	Table 2, p. 80
Bachoo et al. (2013)	AAR	AUS	N	LDE	0.75	150	450	2003-2005	VMB	Disc.	NF	Dummy	Confirm	-0.0157	Table 3, p. 78
Baginski and Rakow (2012)	RAST	US	Y	HDE	1.00	1,355	1,355	2004	VMB	Disc.	FF	SCI	Confirm	-0.1164	Table 4, p. 299
Blanco et al. (2015)	JBFA	US	Y	HDE	1.00	1,667	10,002	2001-2006	VMB	Disc.	PF	SCI	Confirm	-0.1281	Table 6, p. 391
Blanco et al. (2015)	JBFA	US	Y	HDE	1.00	1,667	8,502[b]	2001-2006	REAL	Disc.	PF	SCI	Confirm	-0.0314	Table 10, p. 398
Botosan (1997)	TAR	US	Y	HDE	1.00	122	122	1990	VMB	Disc.	PF	SCI	Confirm	-0.1430	Table 7, p. 342
Botosan and Plumlee (2002)	JAR	US	Y	HDE	1.00	246	2,706	1986-1996	VMB	Disc.	PF	EXI	Mixed	0.0110	Table 4, p. 34
Boujelbene and Affes (2013)	JEFA	FR	N	LDE	0.75	102	102	2009	RFB	Disc.	NF	SCI	Confirm	-0.2180	Table 4, p. 50
Campbell et al. (2014)	RAST	US	Y	HDE	1.00	2,048	8,193	2005-2008	REAL	Disc.	PF	SCI	Confirm	-0.0292	Table 8, p. 436
Cao et al. (2017)	RAST	31 Count.	B	n/a	0.84	6,309	37,856	2004-2009	VMB	Disc.	FF	Dummy	Confirm	-0.1500	Table 2, p. 14
Chen et al. (2003)	SSRN	9 Count.	N	LDE	0.87	273	545	2000-2001	VMB	Disc.	PF	EXI	Confirm	-0.1400	Table 6, p. 391
Cheng et al. (2006)	RQFA	US	Y	HDE	1.00	348	348	2001-2002	VMB	Disc.	FF	EXI	Confirm	-0.0500	Table 2, p. 193
Chien and Lu (2015)	IMDS	US	Y	HDE	1.00	4,122	16,488	2009-2012	RFB	Disc.	NF	SCI	Confirm	-0.0090	Table 4, p. 515
Clarkson et al. (2013)	JAPP	US	Y	HDE	1.00	98	195	2003, 2006	VMB	Disc.	NF	SCI	Reject	0.1096	Table 4. p. 423
Core et al. (2015)	EAR	35 Count.	B	n/a	0.86	3,347	50,201	1990-2004	VMB	Disc.	PF	SCI	Confirm	-0.0083	Table 3, p. 14
Core et al. (2015)	EAR	35 Count.	B	n/a	0.86	3,347	50,201	1990-2004	REAL	Disc.	PF	SCI	Confirm	-0.0030	Table 4, p. 18
Déjean and Martinez (2009)	AIE	FR	N	LDE	0.75	112	112	2006	RFB	Disc.	NF	SCI	Reject	0.1450	Table 5, p. 73
Dhaliwal et al. (2011)	TAR	US	Y	HDE	1.00	795	11,925	1993-2007	VMB	Disc.	NF	Dummy	Mixed	-0.0297	Table 4, p. 76
Dhaliwal et al. (2014)	JAPP	31 Count.	B	n/a	0.83	6,093	79,212	1995-2007	VMB	Disc.	NF	Dummy	Confirm	-0.0700	Table 3, p. 341
Elzahar et al. (2015)	IRFA	UK	N	HDE	0.83	90	448	2006-2010	VMB	Disc.	FF	SCI	Confirm	-0.1695	Table 7, p. 106
Elzahar et al. (2015)	IRFA	UK	N	HDE	0.83	90	448	2006-2010	VMB	Disc.	NF	SCI	Reject	-0.0378	Table 7, p. 106
Embong et al. (2012)	ARA	MY	N	LDE	0.92	135	406	2004-2006	VMB	Disc.	PF	SCI	Mixed	-0.1430	Table 2, p. 126

Table continued next page.

Table 2.4: Information Quantity on CoE – Description of Studies Included in the Meta-Analysis (cont).

Study	Journal[*]	Country[†]	US?[†]	Disc. Env.[€]	Disc. Reg. Score[$]	No. of firms[§]	No. of firm-years[‡]	Sampling Period	DV: CoE[¥]	IV: Quantity[£]	Disc. Type[£]	Disc. Metric[£]	Direct Link?	Effect size (Pearson's *r* coefficient)	Source of Information
Espinosa and Trombetta (2007)	JBFA	SP	N	LDE	0.50	50	250	1998-2002	VMB	Disc.	PF	EXI	Mixed	-0.2000	Table 4, p. 1384
Eugster (2014)	SSRN	CH	N	LDE	0.67	104	1,039	1999-2008	VMB	Disc.	PF	EXI	Mixed	-0.2260	Table 4, p. 40
Evans (2016)	CAR	US	Y	HDE	1.00	187	935	2003-2007	VMB	Disc.	FF	SCI	Mixed	0.0440	Table 1, pp.1147-8
Feng et al. (2009)	SSRN	US	Y	HDE	1.00	335	4,024	1995-2006	VMB	Disc.	FF	SCI	Mixed	-0.0343	Table 4, pp. 30-1
Francis et al. (2005b)	TAR	23 Count.	N	n/a	0.69	137	274	1991, 1993	VMB	Disc.	PF	EXI	Confirm	-0.1450	Table 7, p. 1154
Francis et al. (2008)	JAR	US	Y	HDE	1.00	677	677	2001	VMB	Disc.	PF	SCI	Mixed	-0.0381	Table 4, p. 79
Fu et al. (2012)	JAE	US	Y	HDE	1.00	333	7,654	1951-1973	REAL	Disc.	NF	SCI	Confirm	-0.0600	Table 3, p. 139
Fu et al. (2012)	JAE	US	Y	HDE	1.00	333	7,654	1951-1973	RFB	Disc.	NF	SCI	Confirm	-0.0500	Table 3, p. 139
García-Sánchez and Noguera-Gámez (2017)	IBR	27 Count.	B	n/a	0.72	659	3,294	2009-2013	VMB	Disc.	PF	Dummy	Confirm	-0.0620	Table 5, p. 965
Gietzmann and Ireland (2005)	JBFA	UK	N	HDE	0.83	30	301	1993-2002	VMB	Disc.	PF	SCI	Mixed	-0.1340	Table 2c, p. 625
Grüning (2011)	BR	DE	N	LDE	0.42	361	361	2006	REAL	Disc.	PF	EXI	Confirm	-0.1580	Table 4, p. 58
Hail (2002)	EAR	CH	N	LDE	0.67	73	73	1997	VMB	Disc.	PF	EXI	Confirm	-0.4780	Table 5, p. 761
Healy et al. (1999)	CAR	US	Y	HDE	1.00	37	408	1980-1990	REAL	Disc.	PF	EXI	Confirm	-0.1034	Table 6, p. 504
Khlif et al. (2015)	JAAR	EG	N	LDE	0.50	73	292	2006-2009	RFB	Disc.	PF	SCI	Confirm	-0.2730	Table 4, p. 44
Kim and Shi (2011)	JAPP	US	Y	HDE	1.00	1,066	3,198	2003-2005	VMB	Disc.	FF	Dummy	Mixed	0.0148	Table 5, p. 360
Kothari et al. (2009)	TAR	US	Y	HDE	1.00	223	1,338[a]	1996-2001	RFB	Disc.	PF	SCI	Confirm	-0.0313	Table 3, p. 1658
Kristandl and Bontis (2007)	JIC	4 Count.	N	LDE	0.46	95	95	2004	VMB	Disc.	PF	SCI	Mixed	0.0315	Table 6, p. 587
La Rosa and Liberatore (2014)	EMJ	8 Count.	N	n/a	0.60	62	309	2005-2009	VMB	Disc.	PF	SCI	Reject	0.0560	Table 10, p. 816
Lopes and de Alencar (2010)	IJA	BR	N	LDE	0.25	55	276	1998, 2000/02/04/05	VMB	Disc.	PF	SCI	Confirm	-0.2900	Table 3, p. 454
Mangena et al. (2016)	JAAF	UK	N	HDE	0.83	125	125	2004	VMB	Disc.	FF	SCI	Confirm	-0.2500	Table 3, p. 12
Mangena et al. (2016)	JAAF	UK	N	HDE	0.83	125	125	2004	VMB	Disc.	NF	SCI	Confirm	-0.3440	Table 3, p. 12
Ng and Rezaee (2015)	JCF	Global	B	n/a	n/a	598	13,745	1991-2013	VMB	Disc.	NF	SCI	Confirm	-0.0300	Table 3, pp. 138-9
Ogneva et al. (2007)	TAR	US	Y	HDE	1.00	2,021	2,021	2004	VMB	Disc.	NF	Dummy	Reject	0.0127	Table 2, pp.1268-9

Table continued next page.

Table 2.4: Information Quantity on CoE – Description of Studies Included in the Meta-Analysis (cont.)

Study	Journal*	Country†	US?†	Disc. Env.€	Disc. Reg. Score$	No. of firms§	No. of firm-years‡	Sampling Period	DV: CoE¥	IV: Quantity£	Disc. Type£	Disc. Metric£	Direct Link?	Effect size (Pearson's *r* coefficient)	Source of Information
Orens et al. (2009)	MD	4 Count.	N	LDE	0.55	223	223	2002	VMB	Disc.	NF	SCI	Confirm	-0.1313	Table 4, pp.1547-8
Orens et al. (2010)	JBFA	7 Count.	B	n/a	0.84	668	668	2002	VMB	Disc.	NF	SCI	Confirm	-0.2720	Table 5, p. 1076
Orens et al. (2013)	RAF	4 Count.	N	LDE	0.52	217	217	2002	VMB	Disc.	NF	SCI	Confirm	-0.1150	Table 4, p. 139
Paugam and Ramond (2015)	JBFA	FR	N	LDE	0.75	445	445	2009	VMB	Disc.	FF	SCI	Confirm	-0.1112	Table 4, p. 606
Plumlee et al. (2015)	JAPP	US	Y	HDE	1.00	79	474	2000-2005	VMB	Disc.	NF	SCI	Confirm	-0.0150	Table 3, p. 351
Poshakwale and Courtis (2005)	MDE	Global	B	n/a	n/a	27	135	1995-1999	VMB	Disc.	PF	SCI	Confirm	-0.3410	Table 4, p. 438
Reverte (2012)	CSREM	SP	N	LDE	0.50	19	114	2003-2008	VMB	Disc.	NF	EXI	Confirm	-0.2388	Table 4, p. 263
Richardson and Welker (2001)	AOS	CA	N	HDE	0.92	108	324	1990-1992	VMB	Disc.	FF	EXI	Confirm	-0.0460	Table 2, p. 604
Richardson and Welker (2001)	AOS	CA	N	HDE	0.92	108	324	1990-1992	VMB	Disc.	NF	EXI	Reject	0.0110	Table 2, p. 604
Saini and Herrmann (2012)	SSRN	US	Y	HDE	1.00	87	87	2005	VMB	Disc.	PF	SCI	Confirm	-0.0770	Table 2, p. 42
Tohang and Hutagaol-Martowidjojo (2015)	ASL	ID	N	LDE	0.50	29	58	2010-2011	VMB	Disc.	PF	SCI	Confirm	-0.2210	Table 2, p. 901
Wu et al. (2014)	EMFT	TW	N	LDE	0.75	121	482	2007-2010	VMB	Disc.	NF	Dummy	Confirm	-0.0930	Table 3, p. 113
Xiao-feng et al. (2006)	MSE	CH	N	LDE	0.67	102	102	2005	VMB	Disc.	PF	SCI	Confirm	-0.5200	Table 4, p. 1449
Xu (2009)	GJBR	US	Y	HDE	1.00	212	212	1996	VMB	Disc.	PF	EXI	Mixed	-0.0300	Table 3, p. 21
Zhao et al. (2009)	RAF	US	Y	HDE	1.00	255	255	2000	VMB	Disc.	NF	Dummy	Confirm	-0.1529	Table 9, p. 274

Notes: * *Journal names along with their ABS 2015 ranking are shown in the appendix.* † *AU: Australia; BR: Brazil; CH: Switzerland; CA: Canada: DE: Germany; EG: Egypt; FR: France; ID: Indonesia; MY: Malaysia; SP: Spain; TW: Taiwan; UK: United Kingdom; US: United States; Studies focusing on US firms only are denoted (Y), non-US studies (N) and studies using inseparably both US and non-US firms in their sample (B).* $ *Disc. Reg.: Disclosure regulation score from La Porta et al. (2006); in the case of multi-country studies scores are a weighted average by number of firm-years per country.* § *Number of firms is approximated as number of firm-years divided by number of sample years.* ‡ *When multiple samples are selected for one study, average sample size is reported.* ¥ *REAL: realised-return, RFB: risk factor-based, VMB: valuation model-based cost of equity proxy.* £ *Disc.: Disclosure; Disclosure Types: Full-financial (FF), part-financial (PF), non-financial (NF) disclosure; Disclosure Metrics: Self-constructed index (SCI), external third-party index (EXI), binary dummy variable (Dummy).* a *Converted firm-quarters into firm-years.* b *Converted firm-months into firm-years.*

Table 2.5: Descriptive Statistics of Precision, Asymmetry and Quantity Studies

	Precision		Asymmetry		Quantity		Total	
Sample								
Observations	48	100%	28	100%	62	100%	138	100%
Studies	35	73%	22	79%	56	90%	113	82%
Direct Link								
Accept	27	56%	16	57%	44	71%	87	63%
Reject	8	17%	4	14%	6	10%	18	13%
Mixed	13	27%	8	29%	12	19%	33	24%
Published Work								
Yes	42	88%	25	89%	50	81%	117	85%
No	6	13%	3	11%	12	19%	21	15%
Publ. Quality								
Higher-Tier	29	60%	24	86%	37	60%	90	65%
Lower-Tier	19	40%	4	14%	25	40%	48	35%
Country								
US	41	85%	24	86%	25	40%	90	65%
Non-US	7	15%	4	14%	37	60%	48	35%
CoE Proxy								
REAL	16	33%	21	75%	6	10%	43	31%
VMB	26	54%	7	25%	50	81%	83	60%
RFB	6	13%	0	0%	6	10%	12	9%
Precision Proxy								
Acc. Quality	41	85%					41	30%
Analyst	7	15%					7	5%
Asymmetry Proxy								
Micro.			24	86%			24	17%
Non-Micro.			4	14%			4	3%
Quantity Proxy								
Disclosure					62	100%	62	45%
thereof: FF/PF/NF					11/29/22	18/47/35%	62	45%
thereof: SCI/EXI/Dummy					38/13/11	61/21/18%	62	45%

Notes: This table reports descriptive statistics of the Precision, Asymmetry and Quantity studies included in the literature review for several sample characteristics: (i) Sample: number of studies and observations included; (ii) Direct Link: number of observations (no. of obs.) accepting, rejecting, and finding mixed results for the link with CoE; (iii) Published Work: no. of obs. which are published and unpublished; (iv): Publication Quality: no. of obs. appearing in higher-tier journals (4 & 3 rated journals in ABS 2015 list) and lower-tier journals (ABS 2015 2 & 1 rated, unranked and unpublished work); (v) Country: no. of obs. focusing on US and non-US firms; (vi) CoE Proxy: no. of obs. using realised-return (REAL), risk factor-based (RFB) and valuation model-based (VMB) cost of equity proxies; (vii) Precision Proxy: no. of obs. applying accounting quality (Acc. Quality) and analyst-based (Analyst) proxies; (ix) Asymmetry Proxy: no. of obs. using microstructure (Micro.) and non- microstructure-based proxies; (x): Quantity Proxy: no. of obs. focusing on full-financial (FF), part-financial (PF), non-financial (NF) disclosure and using self-constructed indices (SCI), external third-party indices (EXI) and binary dummy variables (Dummy) to measures quantity.

2.5.1 Information Precision and Cost of Equity

Most of the forty-eight observations analysing the CoE effects of *Precision* relate to published work (n=42; 88%) and appear in higher-tier journals (29; 60%).[10] The sample is strongly tilted towards US firms, with only nine observations (19%) stemming from non-US data. About one-third of observations use realised returns to measure for CoE (REAL: 16; 33%) and the remaining two-thirds mainly apply VMB proxies (VMB: 26; 54%; RFB: 6; 13%). Most observations (41; 85%) rely on accounting/earnings quality measures to proxy for information precision, with analyst-based proxies being the exception (see Table 2.5).

Focusing on the general conclusion of each paper, about half of all observations (n=27; 56%) can be regarded as confirming the negative association between *Precision* and CoE, eight studies tend to reject it (17%), and the remaining papers provide mixed/conditional results for this link (13; 27%). For instance, Ogneva (2012) shows that only after controlling for cash flow shocks in realised returns, a negative association with *Precision* exists; Kim and Qi (2010) confirm H1 after excluding low-priced firms from their sample, and Mashruwala and Mashruwala (2011) document a negative relation only in January. This qualitative assessment is consistent with meta-analytic results (see Table 2.6). The mean effect size between *Precision* and CoE is -0.048 with a 95% confidence interval between -0.142 and 0.047, which illustrates that results for the association between information precision and CoE are mixed across studies. A highly significant χ^2 of 177.89 with 47 degrees of freedom along with only 27 percent of observed variance explained by sampling error (S_e^2/S_r^2) rejects the null of homogenous data and signposts the presence of moderating effects.

2.5.1.1 Measurement of Cost of Equity

Dividing my sample according to CoE measurement yields some interesting insights (Table 2.6, Row II & III). The first sub-group includes those studies which use realised re-

[10] I refer to four and three rated journals (ABS 2015) as higher-tier outlets, and denote two and one rated journals as well as unraked and unpublished work as lower-tier.

turns (REAL) as a proxy for CoE; the second group (RFB/VMB) contains studies applying either risk factor-based or valuation model-based CoE estimates. Total average effect size for the REAL group is -0.014;[11] the one for RFB/VMB is about seven times larger ($\bar{r}$: -0.082) and almost significant at the 10% level (p-value: 0.111). The difference between the two effect sizes is highly significant (t-statistic: -165.1; untabulated two-sample T-test). Overall this is evidence that the measurement of CoE constitutes a moderating effect on results: while there seems to be no relation between *Precision* and firm's realised returns, the association with RFB and VMB proxies is statistically and economically meaningful.[12]

2.5.1.2 Measurement of Information Precision

In a similar vein, the empirical measurement of *Precision* might explain the overall insignificant correlation with CoE ($\bar{r}$: -0.048 ± 0.095).[13] However, irrespective of whether *Precision* is proxied by accounting quality (-0.044 ± 0.085) or analyst forecasts (-0.067 ± 0.139), the associations with CoE remain insignificant (Table 2.6, Row IV). Moreover, results suggest that data heterogeneity stems from RFB/VMB studies, since variation across REAL results is mainly due to sampling error (χ^2: 5.63, df: 13); thus, I focus subsequent analyses on RFB/VMB studies.

Table 2.6 Row II, shows that the relation between CoE and analyst-based proxies is twice as strong ($\bar{r}$: -0.142 vs. -0.074), considerably more significant (p-value: 0.00 vs. 0.122) and less heterogonous (χ^2: 9.03* vs. 83.21***) than for accounting quality-based proxies. This stronger correlation might be explained by the fact that analyst-based proxies are more volatile measures of *Precision* than accounting quality ones, with greater variance leading to higher correlations coefficients. For instance, Barron et al.'s (1998) analyst forecast precision measure is perceived to be highly sensitive to outliers and measurement error (Barron et al., 2012, p. 21), which makes it a noisy and highly dispersed

---

11 As the sampling error variance (Se^2) is larger than the observed variance (Sr^2), the population variance (Sp^2) is negative; thus, no meaningful confidence interval can be determined.

12 The effect size between information precision and RFB proxies ($\bar{r}$: -0.100 ± 0.084) is slightly stronger than for VMB proxies ($\bar{r}$: -0.073 ± 0.103); however, as the number of RFB studies (n=6) is low, a separate sub-group is not meaningful.

13 Average effect size plus/minus two standard deviations (i.e. $\sqrt{Sp^2}$) of the mean (i.e., 95% interval).

proxy for *Precision*. In contrast, Dechow and Dichev's (2002) accrual quality (AQ) metric—the most widely used measure of accounting quality—is estimated from a time-series of firm fundamentals, which increases its robustness and decreases dispersion. To that extent, accounting quality studies draw a more conservative picture of the relation between *Precision* and CoE.

A sufficient large number of accounting quality-based studies (n=41) allows for further analysis of this subset of observations.[14] Designated by a significant chi-square statistic (χ^2: 83.21***, df: 26), the link between accounting quality and RFB/VMB estimates is exposed to moderator effects. As noted before, AQ metrics are a common proxy for accounting quality; Table 2.7, Row II, sub-samples data accordingly and shows that the relation between the AQ-metric group and CoE remains moderated (χ^2: 68.04***, df: 14) and insignificant ($\bar{r}$: -0.082 ± 0.120), while the opposite is observed for *non*-AQ metrics—such as earnings value relevance or accounting conservatism (χ^2: 13.08, df: 11; $\bar{r}$: -0.063 ± 0.020). Differently stated, as the proxies in my sample become more heterogenous, the correlation with CoE becomes less—not more—moderated. This is contrary to expectations in that one would expect that studies sharing the same underlying empirical measure also to find similar results. Therefore, the debate over whether *Precision* impacts CoE is in fact a debate over whether accrual quality affects CoE: results for the remaining studies—which at times use very diverse measures—corroborate the proposition that firms with higher accounting quality enjoy lower CoE.

[14] Given only seven analyst-based studies in my sample, further in-depth analysis for this subset is not meaningful.

Table 2.6: Results by Information Precision and CoE Measures

	Acc.Qual.	**Analyst**	Total
RFB/VMB			
r:	-0.074	-0.142	-0.082
95% CI:	[-0.168; 0.020]	[-0.211; -0.073]	[-0.183; 0.019]
S_e^2/S_r^2:	0.324	0.554	0.304
χ^2_{K-1}:	83.21***	9.03*	105.13***
K:	27	5	32
Sample:	24,194	3,130	27,324
REAL			
r:	-0.014	-0.011	-0.014
95% CI:	[-0.014; -0.014]#	[-0.053; 0.031]	[-0.014; -0.014]#
S_e^2/S_r^2:	2.484	0.517	1.675
χ^2_{K-1}:	5.63	3.87	9.55
K:	14	2	16
Sample:	23,826	4,111	27,937
Total			
r:	-0.044	-0.067	-0.048
95% CI:	[-0.129; 0.040]	[-0.206; 0.071]	[-0.142; 0.047]
S_e^2/S_r^2:	0.313	0.161	0.270
χ^2_{K-1}:	131.18***	43.50***	177.89***
K:	41	7	48
Sample:	48,020	7,241	55,261

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ^2_{K-1}), number of studies (K) and sample size as the average number of firms per year (sample) for the association between Information Precision and CoE. RFB/VMB contains studies using risk factor-based and/or valuation model-based CoE measures; REAL subsumes studies using realised returns as CoE proxy. Distinction is made between studies applying accounting quality (Acc.Qual.) and analyst-based (Analyst) proxies for Precision. # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

Table 2.7: Results by Accounting Quality and CoE Measures

	AQ-Metric	Non-AQ-Metric	Total
RFB/VMB			
r:	-0.082	-0.063	-0.074
95% CI:	[-0.203; 0.038]	[-0.083; -0.043]	[-0.168; 0.020]
S_e^2/S_r^2:	0.220	0.918	0.324
χ^2_{K-1}:	68.04***	13.08	83.21***
K:	15	12	27
Sample:	13,891	10,303	24,194
REAL			
r:	-0.011	-0.020	-0.014
95% CI:	[-0.011; -0.011]#	[-0.020; -0.020]#	[-0.014; -0.014]#
S_e^2/S_r^2:	2.060	3.889	1.675
χ^2_{K-1}:	3.40	1.80	5.63
K:	7	7	14
Sample:	13,998	9,827	23,826
Total			
r:	-0.046	-0.042	-0.044
95% CI:	[-0.154; 0.061]	[-0.074; -0.010]	[-0.129; 0.040]
S_e^2/S_r^2:	0.206	0.780	0.313
χ^2_{K-1}:	106.59***	24.37	131.18***
K:	22	19	41
Sample:	27,889	20,131	48,020

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ^2_{K-1}), number of studies (K) and sample size as the average number of firms per year (sample) for the association between Accounting Quality and CoE. RFB/VMB contains studies using risk factor-based and/or valuation model-based CoE measures; REAL subsumes studies using realised returns as CoE proxy. Distinction is made between studies applying an accrual quality metric (AQ-Metric) and those alternative accounting quality measures (Non-AQ-Metric). # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

2.5.1.3 Publication Bias

To address concerns of publication bias (Møller and Jennions, 2001), I analyse if there are differences in result between higher-tier journal and lower-tier journals. As shown in Table 2.8, differences in average effect sizes between the two groups follow no clear pattern. For instance, when measuring CoE by RFB/VMB proxies, total average affect size is larger in higher-tier journals ($\bar{r}$: -0.097) than in lower-tier ones (-0.067); when focusing on realised return-based studies, total average effect size is larger for lower-tier

(-0.019) vis-à-vis higher-tier publications (-0.014). However, results in higher-tier journals are slightly more heterogenous than in lower-tier journals, which can be inferred from more significant chi-squares, lower S_e^2/S_r^2 ratios and wider confidence intervals. Therefore, results tend to be free of publication bias, but the relationship between *Precision* and CoE is discussed more controversial in higher-tier outlets.

Table 2.8: Results by Accounting Quality, CoE Measures and Publication Quality

	Acc.Qual.		**Analyst**		Total	
RFB/VMB	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-0.092	-0.054	-0.154	-0.136	-0.097	-0.067
95% CI:	[-0.194; 0.009]	[-0.119; 0.011]	[-0.248; -0.060]	[-0.187; -0.086]	[-0.202; 0.009]	[-0.153; 0.019]
S_e^2/S_r^2:	0.273	0.523	0.555	0.580	0.283	0.376
χ_{K-1}^2:	47.6***	26.8**	5.4*	3.4*	56.5***	42.6***
K:	13	14	3	2	16	16
Sample:	12,698	11,496	994	2,136	13,692	13,632
REAL	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-0.014	-0.019	-0.011	-	-0.014	-0.019
95% CI:	[-0.014; -0.014]#	[-0.019; -0.019#	[-0.053; 0.031]	-	[-0.014; -0.014]#	[-0.019; -0.019#
S_e^2/S_r^2:	2.019	20.503	0.517	-	1.389	20.503
χ_{K-1}^2:	5.4	0.1	3.9	-	9.4	0.1
K:	11	3	2	-	13	3
Sample:	22,081	1,745	4,111	-	26,192	1,745
Total	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-0.043	-0.049	-0.039	-0.136	-0.042	-0.061
95% CI:	[-0.135; 0.050]	[-0.108; 0.009]	[-0.162; 0.085]	[-0.187; 0.086]	[-0.139; 0.055]	[-0.144; 0.021]
S_e^2/S_r^2:	0.235	0.591	0.197	0.580	0.283	0.411
χ_{K-1}^2:	102.0***	28.8**	25.4***	3.4*	127.5***	46.3***
K:	24	17	5	2	29	19
Sample:	34,779	13,241	5,105	2,136	39,884	15,377

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ_{K-1}^2), number of studies (K) and sample size as the average number of firms per year (sample) for the association between Information Precision and CoE. RFB/VMB contains studies using risk factor-based and/or valuation model-based CoE measures; REAL subsumes studies using realised returns as CoE proxy. Distinction is made between studies applying accounting quality (Acc.Qual.) and analyst-based (Analyst) proxies for Precision. Higher-Tier: 4 & 3 rated journals in ABS2015; Lower-Tier: 2 & 1 rated journals in ABS 2015, unranked and unpublished work. # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

Overall, results suggest that the relationship between information precision and CoE depends on the measurement of CoE: higher *Precision* leads to lower CoE if measured by RFB/VMB proxies, but this link is trivial in asset pricing tests and when future realised returns are used. Furthermore, a material association between RFB/VMB proxies and analyst-based as well as *non*-AQ metric-based studies is found; however, this relation is insignificant in studies using AQ metrics as measures of accounting quality. This warrants the conclusion that the controversy over the impact of information precision on firms' CoE stems, by and large, from the debate on the market pricing of accrual quality.

2.5.2 Information Asymmetry and Cost of Equity

Twenty-eight observations in my sample examine the link between *Asymmetry* and CoE, with most observations relating to published work (n=25; 89%) and appearing in higher-tier journals (24; 86%). The sample is strongly tilted towards US studies, with only four observations (14%) relying on non-US data. Three-quarters of observations apply realised returns to proxy for CoE (REAL: 21; 75%) and the remainder exclusively use VMB estimates (VMB: 7; 25%). The vast majority of observations (24; 86%) utilises market microstructure proxies (i.e., PIN scores, bid-ask-spreads), whereas non-microstructure estimates (e.g., analyst-based proxies) are rarely used (see Table 2.5).

As noted above, a meaningful meta-analysis for the link between *Asymmetry* and CoE is not possible; therefore, proceeding analyses focus on the general conclusion of each paper. Figure 2.2, Panel A, shows that about half of all observations confirm the positive association between *Asymmetry* and CoE (n=16; 57%), four observations (14%) tend to reject it, and eight observations (29%) provide mixed/conditional results. In particular, the level of market competition seems an important conditioning variable, given that *Asymmetry* effects tend to vanish as markets become perfectly competitive (Akins et al. (2012), Armstrong et al. (2011), Luong et al. (2011)). Furthermore, findings in Eleswarapu and Reinganum (1993) and Kang (2010) indicate the existence of January-effects.

2.5.2.1 Measurement of Cost of Equity

Sub-sampling observations according to CoE measurement reveals some interesting points (see Figure 2.2, Panel B and C). First, the literature mainly uses realised returns to proxy for firms' CoE (REAL: 21; 75%), with seven observations (25%) relying on VMB estimates, and none using traditional RFB ones. Second—and more importantly—there is notable variation in the conclusions reached by REAL studies (reject: 14%; mixed: 38%; confirm: 48%), but findings are rather uniform when VMB proxies are used (reject: 14%; confirm: 86%).

Given that realised returns tend to be noisy proxies of expected returns (Chen et al. (2013), Fama and French, 1997, Vuolteenaho (2002)), one might argue that greater weight should be placed on studies applying VMB (i.e., ICC) estimates. In other words, the reason why "only" 48% percent of REAL studies confirm the positive association between *Asymmetry* and CoE might stem from imprecise CoE estimates. While convincing evidence exists that VMB proxies show greater validity than realised return-based proxies (Botosan and Plumlee (2005), Botosan et al. (2011), Lee et al. (2010, 2015)), those estimates are not impeccable either. Beyond the issue of lacking estimates for young, small and financial distressed firms due to coverage bias by analysts (Diether et al. (2002), Hong et al. (2000), La Porta (1996)), it is in particular the problem of upward biased ICC estimates—due to optimistic analyst forecasts (Dugar and Nathan (1995), Francis and Philbrick (1993), McNichols and O'Brien (1997))—which raises concerns. For example, Hwang et al. (2013) conclude that "as long as the [ICC] estimates are derived from analysts' earnings forecasts, potential measurement errors in [ICC] estimates could remain and influence […] findings"; therefore, the authors call for "more efforts to fine-tune [ICC] measures" (p. 165).

Overall, findings suggest that researchers' choice of CoE measurement affects the conclusions reached regarding the relation between *Asymmetry* and CoE, with extant work facing a trade-off between either using comprehensive samples (with noisy return estimates) or valid CoE proxies (with biased samples). This makes it difficult to disentangle the economic impact from the impact attributable to measurement errors and sampling bias when analysing this link; however, recent advancements in model-based ICC research might help to overcome this predicament in future work (Hou et al. (2012), Li and Mohanram (2014)).

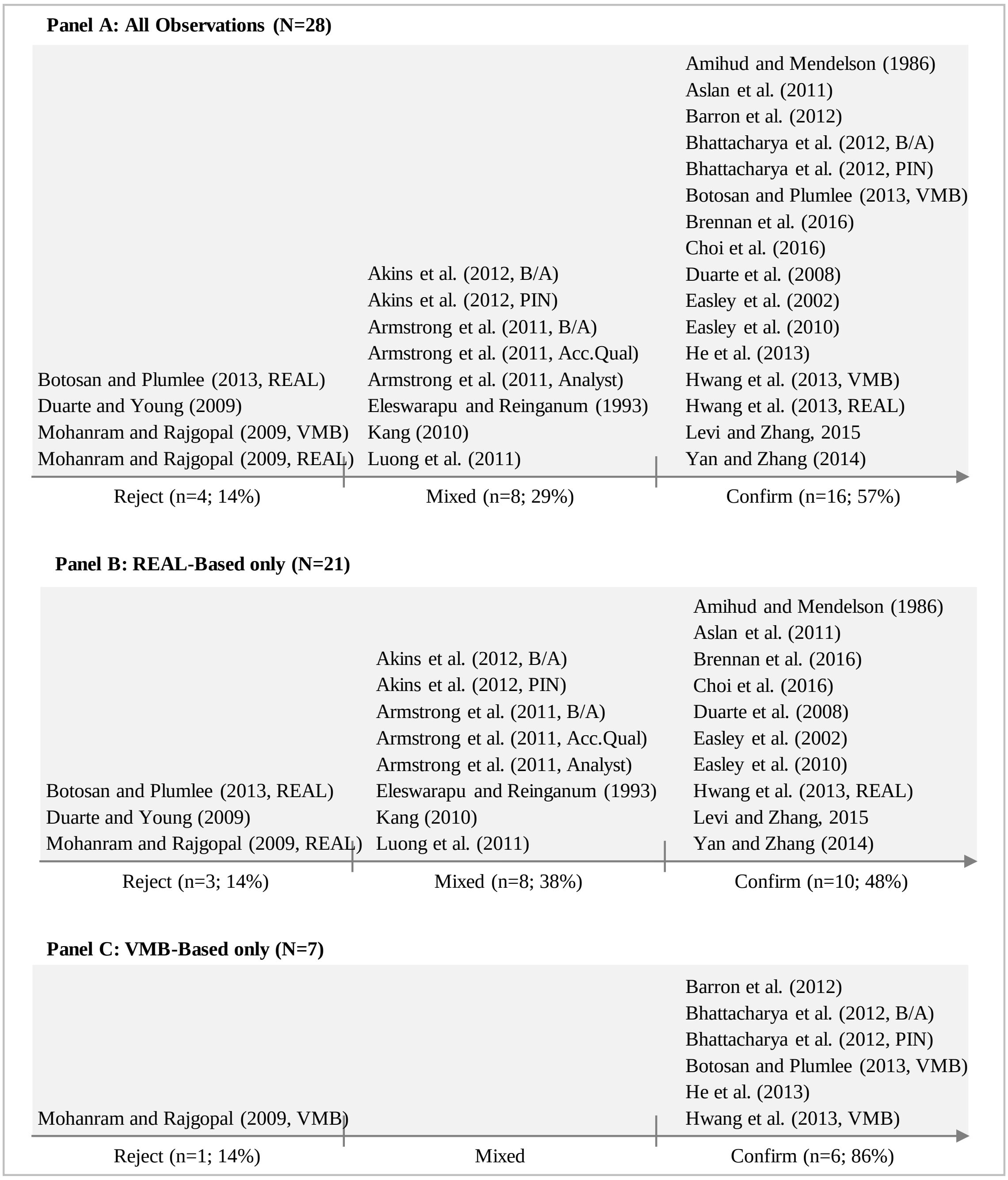


Figure 2.2: Information Asymmetry on CoE – Measurement of Cost of Equity

The figure categorises the sample observations in respect to their general conclusion reached regarding Hypothesis 2. Panel A shows results for all observations. Panel B and Panel C for REAL-based and VMB-Based observations, respectively.

2.5.2.2 Measurement of Information Asymmetry

As most studies in my sample rely on market microstructure-based proxies to measure information asymmetry (n=24; 86%), I concentrate subsequent analyses on this subset of observations. Figure 2.3 shows that PIN scores are most widely used (17; 71%) in the literature, followed by bid/ask spreads operated in some studies (7; 29%). None of the spread-based studies reject the direct link between *Asymmetry* and CoE (Figure 2.3, Panel C), which suggests that higher bid/ask spreads indicate greater information asymmetry and can induce higher CoE. In contrast, the association between PIN scores and expected rate of returns is somewhat debated (reject: 24%, mixed: 18%, confirm: 59%): Mohanram and Rajgopal (2009, p. 241) conclude that "there is not much evidence [that] PIN is a source of priced information risk" and Duarte and Young (2009) show that it is the illiquidity component of PIN which explains the positive relation with CoE, and not the *Asymmetry* part of it. However, the majority of studies confirm the negative CoE effects arising from asymmetric information (Figure 2.3, Panel B); together with new evidence from longer sampling periods (e.g., Aslan et al. (2011), Brennan et al. (2016)) and improved estimation techniques (Hwang et al., 2013), this substantiates the proposition of PIN being an important driver of CoE.[15]

Taken findings together, prior evidence corroborates the conjecture of informational disparity between investor groups causing an increase in firms' CoE. However, recent analytical and empirical evidence demonstrates that high levels of market competition tend to subdue these *Asymmetry* effects. More specifically, Akins et al. (2012), Armstrong et al. (2011) and Luong et al. (2011) show that—consistent with the analytical model of Lambert et al. (2012)—the significance of *Asymmetry* on CoE declines as market competition increases; that is, in perfectly liquid markets, in which both informed and uninformed investors act as price takers, asymmetric information has no material effect on firms' CoE. This puts forward market competition as an important conditioning variable in future research.[16]

---

[15] See also Boehmer et al. (2007), Yan and Zhang (2012) and William Lin and Ke (2011) for a discussion on how to improve PIN score estimates.

[16] Common proxies for market competition include: investor concentration (Akins et al., 2012), institutional ownership (Luong et al., 2011) and number of shareholders (Armstrong et al. (2011), Barron et al. (2012)).

Panel A: All Microstructure-Based Observations (N=24)

Reject (n=4; 17%)	Mixed (n=6; 25%)	Confirm (n=14; 58%)
Botosan and Plumlee (2013, REAL)	Akins et al. (2012, B/A)	Amihud and Mendelson (1986)
Duarte and Young (2009)	Akins et al. (2012, PIN)	Aslan et al. (2011)
Mohanram and Rajgopal (2009, VMB)	Armstrong et al. (2011, B/A)	Bhattacharya et al. (2012, B/A)
Mohanram and Rajgopal (2009, REAL)	Eleswarapu and Reinganum (1993)	Bhattacharya et al. (2012, PIN)
	Kang (2010)	Botosan and Plumlee (2013, VMB)
	Luong et al. (2011)	Brennan et al. (2016)
		Duarte et al. (2008)
		Easley et al. (2002)
		Easley et al. (2010)
		He et al. (2013)
		Hwang et al. (2013, VMB)
		Hwang et al. (2013, REAL)
		Levi and Zhang, 2015
		Yan and Zhang (2014)

Panel B: PIN (N=17)

Reject (n=4; 24%)	Mixed (n=3; 18%)	Confirm (n=10; 59%)
Botosan and Plumlee (2013, REAL)	Akins et al. (2012, PIN)	Aslan et al. (2011)
Duarte and Young (2009)	Kang (2010)	Bhattacharya et al. (2012, PIN)
Mohanram and Rajgopal (2009, VMB)	Luong et al. (2011)	Botosan and Plumlee (2013, VMB)
Mohanram and Rajgopal (2009, REAL)		Brennan et al. (2016)
		Duarte et al. (2008)
		Easley et al. (2002)
		Easley et al. (2010)
		Hwang et al. (2013, VMB)
		Hwang et al. (2013, REAL)
		Yan and Zhang (2014)

Panel C: Bid-Ask Spreads (N=7)

Reject	Mixed (n=3; 43%)	Confirm (n=4; 57%)
	Akins et al. (2012, B/A)	Amihud and Mendelson (1986)
	Armstrong et al. (2011, B/A)	Bhattacharya et al. (2012, B/A)
	Eleswarapu and Reinganum (1993)	He et al. (2013)
		Levi and Zhang, 2015

Figure 2.3: Information Asymmetry on CoE – Measurement of Information Asymmetry

The figure categorises the sample observations in respect to their general conclusion reached regarding Hypothesis 2. Panel A shows results for all microstructure-based observations. Panel B and Panel C for PIN-based and Bid/Ask-Spread-Based observations, respectively.

2.5.3 Information Quantity and Cost of Equity

Subsequent meta-analysis is limited to the disclosure literature, given that "simple-proxy" studies are sparse (Barry and Brown (1984, 1985), Clarkson and Thompson (1990), Kumar et al. (2008), Zhang (2006)).[17] Sixty-two observations in my sample analyse the link between *Quantity* and CoE of which most work pertain to published work (n=50; 81%) and appear in higher-tier journals (37; 60%). Ten percent of studies use realised returns to proxy for expected returns (n=6), with VMB measures prevailing in the disclosure literature (VMB: 50; 81%; RFB: 6; 10%). Whilst some studies include firms from multiple countries in their sample (13; 21%), the great majority are single country studies (49; 79%). Interestingly, only twenty-five observations (40%) relate to US firms—which is exceptionally low compared to the US-bias in *Precision* and *Asymmetry* studies (81 and 86 percent, respectively)—and the remaining thirty-seven studies either focus on Non-US countries (29; 47%) or inseparably include both US and Non-US firms (8; 13%) in their samples (see Table 2.5).

Table 2.5 also reports the type of disclosure examined and the disclosure metric used by researchers. Eleven observations (18%) are categorised as having a clear focus on financial disclosure, twenty-two (35%) as evaluating non-financial disclosure aspects, and the remining twenty-nine (47%) as partial-financial (i.e. all those studies with an emphasis on the general quality of firms' disclosure). Most studies apply self-constructed indices to measure disclosure levels (38; 61%), followed by thirteen studies (21%) relying on third-party providers (e.g., AIMR, S&P scores) and eleven (18%) observations using simple dummy variables to distinguish between disclosing and non-disclosing firms.

Examining the general conclusion of each paper, about seventy percent of all observations (n=44; 71%) tend to confirm the negative association between *Quantity* and CoE, six studies reject it (10%), and the remaining twelve papers (19%) provide mixed/conditional results for this link. For instance, Francis et al. (2008) show that the negative relation between disclosure and expected returns tend to vanish after controlling for earnings quality; Evans (2016) and Kim and Shi (2011) suggest that the timeliness and the sign of earnings announcements (good/bad) as well as the degree of market competition are important conditioning variables; and Espinosa and Trombetta (2007) and Gietzmann and

[17] Recently, Souissi and Khlif (2012) also meta-analyse the impact of disclosure on cost of equity capital; however, my analysis differs from theirs in that it operates a larger sample (56 vs. 22 studies), covers a longer sampling period (1997-2010 vs. 1997-2017) and analyses substantially more firm-years (342,116 vs 9,553). Furthermore, I include both mandatory and voluntary disclosure studies, while Souissi and Khlif (2012) focus on voluntary disclosure only.

Ireland (2005) demonstrate that the impact of disclosure on CoE is extenuated by accounting conservatism (i.e., only firms adopting aggressive accounting policies can reduce their CoE by increased disclosure activity). This qualitative assessment is consistent with meta-analytic results (see Table 2.9). The mean effect size between *Quantity* and CoE is -0.066 with a 95% confidence interval between -0.190 and 0.058, illustrating that findings vary across studies. A highly significant χ^2 of 241.21 with 61 degrees of freedom along with only 26 percent of observed variance explained by sampling error (S_e^2/S_r^2) rejects the null of homogenous data and pinpoints the presence of moderating effects.

2.5.3.1 Measurement of Cost of Equity

As before, I divide my sample into two sub-groups according to CoE measurement (Table 2.9, Row II & III). Total average effect size for the RFB/VMB group ($\bar{r}$: -0.075 ± 0.129) is about three times larger than for the REAL group (-0.026 ± 0.037), however, none of the REAL studies examines the association in a pure financial disclosure setting, where CoE effects are most pronounced, which tend to explain this difference in effect size (Columns II-IV).[18] More importantly, both effect sizes are insignificantly different from zero (p-value: 0.256 and 0.171), which indicates that the choice of CoE measurement does *not* explain mixed results in the disclosure literature.

2.5.3.2 Measurement of Quantity

Table 2.9 distinguishes studies according to which type of disclosure is examined by researchers: financial (FF), partial-financial (PF) and non-financial (NF) disclosure studies. Row IV shows an economically and statistically significant correlation of about 12 percent between CoE and FF studies (p-value: 0.019), but a markedly reduced and statistically insignificant effect size of only five percent for PF and NF studies (p-values: 0.358 and 0.361). This reveals that financial disclosure is twice as important to investors than non-financial and partial-financial disclosure. However, this is not to say that non-financial information is irrelevant; for example, a relatively large strand of research within the

[18] The effect size for VMB proxies ($\bar{r}$: -0.084 ± 0.128) is markedly stronger than for RFB measures ($\bar{r}$: -0.017 ± 0.071) which—as in the case of the REAL group—is attributable to the fact that none of the RFB studies focus on financial disclosure.

NF category (n=10; 16%) documents a significantly inverse association between Corporate Social Responsibility disclosure and CoE ($\bar{r}$: -0.056 ± 0.022, unreported).[19] Moreover, findings are robust to researchers' choice of how to measure disclosure levels; that is, irrespective of using self-constructed disclosure indices (SCI) or simple dummy variables to proxy for disclosing and non-disclosing firms, the effect size is always significantly negative for FF studies (SCI: -0.101 ± 0.069; Dummy: -0.126± 0.106), but insignificant and much weaker for NF and PF studies (see Table 2.10).[20]

Next, I examine if different disclosure requirements across countries moderate results. In doing so, I categorise the studies in three different ways: first, I distinguish between US and Non-US studies; second, I assign each study a disclosure regulation score and allocate studies with a score below the sample average of 0.83 to the LOW group and the remainder to the HIGH group;[21] third, I follow Souissi and Khlif (2012) and form groups based on countries transparency culture, where the high disclosure environment group (HDE) includes US, UK and Canadian studies and the low disclosure environment group (LDE) covers the remaining countries in my sample.[22]

Table 2.11, Column V, shows that total disclosure effects are about 3.5 times larger in Non-US ($\bar{r}$: -0.139) than in US studies (-0.039). Similarly, studies concentrating on less regulated (LOW: -0.126) and transparent countries (LDE: -0.141) document an approximately three times stronger correlation between disclosure and CoE than studies focusing on more regulated and transparent countries (HIGH: -0.062; HDE: -0.043). However, the magnitude of these differences varies by disclosure types; for instance, the CoE effect of financial disclosure in US settings is "only" 2.5 times larger than in Non-US settings,

[19] Corporate Social Responsibility studies: Bachoo et al. (2013), Clarkson et al. (2013), Déjean and Martinez (2009), Dhaliwal et al. (2011), Dhaliwal et al. (2014), Ng and Rezaee (2015), Plumlee et al. (2015), Reverte (2012), Richardson and Welker (2001), Wu et al. (2014).

[20] Given that error variances (*Se^2*) for FF and NF studies using externally provided disclosures scores (EXI) are larger than observed variances (*Sr^2*), confidence intervals cannot be calculated and a comparison with PF studies is not meaningful.

[21] Consistent with Core et al. (2015), I measure the level of disclosure regulation by the index of disclosure requirements in securities offerings from La Porta et al. (2006). In the case of multi-country studies, I report a weighted average per observation (weight: firm-years per country).

[22] LDE countries: Austria, Bahrain, Belgium, Brazil, Denmark, Egypt, France, Germany, Hong Kong, India, Indonesia, Korea, Kuwait, Malaysia, Netherlands, Oman, Philippines, Qatar, Saudi Arabia, Singapore, Spain, Sweden Switzerland, Taiwan, United Arab Emirates.

while partial- and non-financial disclosure effects are about three and 4.5 times larger (similar patterns can be observed for HIGH and LOW as well as HDE and LDE studies). What is more, within each sub-sample a substantial amount of variation in results is now attributable to sampling error (e.g., S_e^2/S_r^2 ratios of 0.54 and 0.48 for US and Non-US studies vis-à-vis 0.26 for the sample as a whole); together with less significant chi-square statistics, this confirms that disclosure environments across countries moderate results.

Overall, results confirm that the more financial information firms disclose, the greater their CoE benefits. However, findings also suggest that in countries where disclosure regulation and requirements are strong (such as the US), investors appreciate firms' disclosure efforts to a much lesser extent than in weaker regulated countries.

2.5.3.3 Publication Bias

As before, I analyse if there are differences in result between higher-tier and lower-tier publications. Table 2.12 shows that differences in average effect sizes between the two groups follow no clear pattern: higher-tier journals report a stronger effect size when RFB/VMB proxies are used (high: -0.081; low: -0.055), while lower-tier journals' effect size is stronger for REAL proxies (-0.020; -0.158); similarly, higher-tier studies find stronger CoE effects for financial disclosure (-0.119; -0.044), while lower-tier studies report a stronger relation with partial-financial disclosure (-0.038; -0.138)—average effect sizes are similar for non-financial disclosure (-0.061; -0.040). However, results in higher-tier journals are somewhat more heterogenous than in lower-tier journals as indicated by more significant chi-squares, lower $\mathrm{S_e^2/S_r^2}$ ratios and wider confidence intervals. This suggests that results are free of publication bias, but the impact of corporate disclosure on firms' CoE is debated more controversially in higher-tier outlets.

To summarize, findings show that researchers' choice of CoE measurement does *not* explain mixed results in the disclosure literature, but it is the type of disclosure (financial vs. partial-/non-financial) as well as the disclosure setting (strongly vs. weakly regulated countries) which moderates results. Robustness tests show that these findings are widely insensitive to how disclosure levels are measured by researchers.

Table 2.9: Results by Disclosure Types and CoE Measures

	Full-Financial (FF)	Partial-Financial (PF)	Non-Financial (NF)	Total
RFB/VMB				
r:	-0.117	-0.069	-0.053	-0.075
95% CI:	[-0.215; -0.018]	[-0.215; 0.078]	[-0.167; 0.060]	[-0.204; 0.054]
S_e^2/S_r^2:	0.288	0.331	0.263	0.258
χ^2_{K-1}:	38.14***	72.55***	79.85***	217.38***
K:	11	24	21	56
Sample:	10,423	8,625	17,560	36,607
REAL				
r:	-	-0.025	-0.060	-0.026
95% CI:	-	[-0.066; 0.017]	[-0.060; -0.060]#	[-0.063; 0.011]
S_e^2/S_r^2:	-	0.599	-	0.686
χ^2_{K-1}:	-	8.34	-	8.75
K:	-	5	1	6
Sample:	-	7,460	333	7,793
Total				
r:	-0.117	-0.048	-0.054	-0.066
95% CI:	[-0.215; -0.018]	[-0.167; 0.071]	[-0.165; 0.058]	[-0.190; 0.058]
S_e^2/S_r^2:	0.288	0.328	0.275	0.257
χ^2_{K-1}:	38.14***	88.42***	79.86***	241.21***
K:	11	29	22	62
Sample:	10,423	16,085	17,892	44,400

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ^2_{K-1}), number of studies (K) and sample size as the average number of firms per year (sample) for the association between information quantity and CoE. RFB/VMB contains studies using risk factor-based and/or valuation model-based CoE measures; REAL subsumes studies using realised returns as CoE proxy. Distinction is made between studies focusing fully (FF), partially (PF) or not at all (NF) on financial disclosure. # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

Table 2.10: Results by Disclosure Types and Disclosure Metrics

	Full-Financial (FF)	**Partial-Financial (PF)**	**Non-Financial (NF)**	Total
SCI				
r:	-0.101	-0.039	-0.052	-0.050
95% CI:	[-0.215; -0.018]	[-0.145; 0.067]	[-0.217; 0.12]	[-0.179; 0.079]
S_e^2/S_r^2:	0.654	0.317	0.205	0.270
χ^2_{K-1}:	9.17*	60.00***	63.36***	140.89***
K:	6	19	13	38
Sample:	2,537	13,934	7,098	23,569
EXI				
r:	-0.049	-0.128	-0.026	-0.104
95% CI:	[-0.049; -0.049]#	[-0.278; 0.022]	[-0.026; -0.026]#	[-0.231; 0.023]
S_e^2/S_r^2:	1509.313	0.500	1.980	0.594
χ^2_{K-1}:	0.00	18.00**	1.01	21.89**
K:	2	9	2	13
Sample:	456	1,492	127	2,075
Dummy				
r:	-0.126	-0.062	-0.055	-0.083
95% CI:	[-0.232; -0.020]	[-0.062; -0.062]#	[-0.109; 0.00]	[-0.186; 0.020]
S_e^2/S_r^2:	0.117	-	0.456	0.173
χ^2_{K-1}:	25.60***	-	15.36**	63.47***
K:	3	1	7	11
Sample:	7,430	659	10,667	18,756
Total				
r:	-0.117	-0.048	-0.054	-0.066
95% CI:	[-0.215; -0.018]	[-0.167; 0.071]	[-0.165; 0.058]	[-0.190; 0.058]
S_e^2/S_r^2:	0.288	0.328	0.275	0.257
χ^2_{K-1}:	38.14***	88.42***	79.86***	241.21***
K:	11	29	22	62
Sample:	10,423	16,085	17,892	44,400

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ^2_{K-1}), number of studies (K) and sample size as the average number of firms per year (sample) for the association between information quantity and CoE. Distinction is made between type of disclosures—studies focusing fully (FF), partially (PF) or not at all (NF) on financial disclosure—and disclosure metrics— studies using self-constructed disclosure indexes (SCI), externally provided third-party indices (EXI) or simple dummy variables to distinguish between disclosing and non-disclosing firms (Dummy). # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

Table 2.11: Results by Disclosure Types and Disclosure Regimes

Panel A: US and Non-US Firms				
	Full-Financial (FF)	**Partial-Financial (PF)**	**Non-Financial (NF)**	Total
US				
r:	-0.050	-0.056	-0.022	-0.039
95% CI:	[-0.136; 0.035]	[-0.102; -0.010]	[-0.073; 0.028]	[-0.104; 0.026]
S_e^2/S_r^2:	0.484	0.720	0.592	0.539
χ^2_{K-1}:	12.39*	13.88	15.20*	46.38***
K:	6	10	9	25
Sample:	3,346	6,986	9,267	19,599
Non-US				
r:	-0.131	-0.173	-0.098	-0.139
95% CI:	[-0.131; -0.131]#	[-0.375; 0.028]	[-0.264; 0.069]	[-0.361; 0.037]
S_e^2/S_r^2:	1.352	0.437	0.517	0.478
χ^2_{K-1}:	2.96	34.33***	19.33**	60.68***
K:	4	15	10	29
Sample:	768	1,720	1,266	3,754
Panel B: High and Low Disclosure Regulation				
HIGH				
r:	-0.117	-0.035	-0.053	-0.062
95% CI:	[-0.220; -0.014]	[-0.093; 0.024]	[-0.166; 0.061]	[-0.176; 0.052]
S_e^2/S_r^2:	0.262	0.544	0.203	0.221
χ^2_{K-1}:	38.13***	27.58**	69.05***	176.65***
K:	10	15	14	39
Sample:	9,978	14,118	16,351	40,446
LOW				
r:	-0.111	-0.153	-0.083	-0.126
95% CI:	[-0.111; -0.111]#	[-0.381; 0.075]	[-0.187; 0.021]	[-0.314; 0.061]
S_e^2/S_r^2:	-	0.320	0.721	0.400
χ^2_{K-1}:	-	37.47***	9.71	50.04***
K:	1	12	7	20
Sample:	445	1,800	944	3,188

Table continued next page.

Table 2.11: Results by Disclosure Types and Disclosure Regimes (cont.)

Panel C: High (HDE) and Low Disclosure Environments (LDE)				
	Full-Financial (FF)	**Partial-Financial (PF)**	**Non-Financial (NF)**	Total
HDE				
r:	-0.060	-0.056	-0.026	-0.043
95% CI:	[-0.159; 0.039]	[-0.097; -0.015]	[-0.107; 0.054]	[-0.122; 0.037]
S_e^2/S_r^2:	0.490	0.782	0.426	0.489
χ_{K-1}^2:	18.38**	14.07	28.14***	65.46***
K:	9	11	12	32
Sample:	3,668	7,016	9,590	20,275
LDE				
r:	-0.111	-0.186	-0.083	-0.141
95% CI:	[-0.111; -0.111]#	[-0.398; 0.026]	[-0.187; 0.021]	[-0.326; 0.044]
S_e^2/S_r^2:	-	0.390	0.721	0.429
χ_{K-1}^2:	-	30.74***	9.71	46.66***
K:	1	12	7	20
Sample:	445	1,491	944	2,879
Total				
r:	-0.117	-0.048	-0.054	-0.066
95% CI:	[-0.215; -0.018]	[-0.167; 0.071]	[-0.165; 0.058]	[-0.190; 0.058]
S_e^2/S_r^2:	0.288	0.328	0.275	0.257
χ_{K-1}^2:	38.14***	88.42***	79.86***	241.21***
K:	11	29	22	62
Sample:	10,423	16,085	17,892	44,400

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ_{K-1}^2), number of studies (K) and sample size as the average number of firms per year (sample) for the association between information quantity and CoE. Distinction is made between types of disclosures (i.e., studies focusing fully (FF), partially (PF) or not at all (NF) on financial disclosure) and the following disclosure regimes: Panel A reports results for US and non-US firms, Panel B categorises studies according to their disclosure regulation scores, with studies below the sample average of 0.83 being assigned to the LOW group and the remainder to the HIGH group; Panel C distinguishes between high disclosure environments (HDE: US, CA, UK) and a low disclosure environments (LDE: AUS, BR, CH, DE, EG, FR, ID, MY, SP, TW) with the respective countries assigned to each group in parentheses. # zero residual variance is used for CI calculation, given the error variance (Se^2) being larger than the observed variance (Sr^2) resulting in a negative population variance (Sp^2). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

Table 2.12: Results by Disclosure Types, CoE Measures and Publication Quality

	Full-Financial (FF)		**Partial-Financial (PF)**		**Non-Financial (NF)**		Total	
RFB/VMB	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-0.119	-0.044	-0.057	-0.133	-0.061	-0.040	-0.081	-0.055
95% CI:	[-0.219; -0.020]	[-0.044; -0.044]	[-0.172; 0.058]	[-0.352; 0.085]	[-0.187; 0.064]	[-0.121; 0.041]	[-0.208; 0.047]	[-0.184; 0.073]
S_e^2/S_r^2:	0.252	9.654	0.322	0.410	0.194	0.469	0.209	0.401
χ_{K-1}^2:	35.8***	0.2	37.3***	29.3***	56.7***	21.3**	153.3***	59.8***
K:	9	2	12	12	11	10	32	24
Sample:	10,033	390	7,286	1,340	11,007	6,553	28,325	8,282
REAL	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-	-	-0.018	-0.158	-0.060	-	-0.020	-0.158
95% CI:	-	-	[-0.018; -0.018]#	[-0.158; -0.158]#	[-0.060; -0.060]#	-	[-0.020; -0.020]#	[-0.158; -0.158]#
S_e^2/S_r^2:	-	-	2.534	-	-	-	2.329	-
χ_{K-1}^2:	-	-	1.6	-	-	-	2.2	-
K:	-	-	4	1	1	-	5	1
Sample:	-	-	7,099	361	333	-	7,432	361
Total	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier	Higher-Tier	Lower-Tier
r:	-0.119	-0.044	-0.038	-0.138	-0.061	-0.040	-0.068	-0.059
95% CI:	[-0.219; -0.020]	[-0.044; -0.044]#	[-0.124; 0.049]	[-0.328; 0.051]	[-0.184; 0.061]	[-0.121; 0.041]	[-0.170; 0.034]	[-0.190; 0.071]
S_e^2/S_r^2:	0.252	9.654	0.362	0.440	0.212	0.469	0.209	0.393
χ_{K-1}^2:	35.8***	0.2	44.2***	29.6***	56.7***	21.3**	177.1***	63.6***
K:	9	2	16	13	12	10	37	25
Sample:	10,033	390	14,385	1,701	11,340	6,553	35,757	8,643

*Notes: This contingency table reports average effect size (r), the 95 percent confidence interval (95% CI), the sampling-error explained percentage of observed variance (S_e^2/S_r^2), the chi-square statistic (χ_{K-1}^2), number of studies (K) and sample size as the average number of firms per year (sample) for the association between information quantity and CoE. RFB/VMB contains studies using risk factor-based and/or valuation model-based CoE measures; REAL subsumes studies using realised returns as CoE proxy. Distinction is made between studies focusing fully (FF), partially (PF) or not at all (NF) on financial disclosure. Higher-Tier: 4 & 3 rated journals in ABS2015; Lower-Tier: 2 & 1 rated journals in ABS 2015, unranked and unpublished work. # zero residual variance is used for CI calculation, given the error variance (Se²) being larger than the observed variance (Sr²) resulting in a negative population variance (Sp²). ***, **, * denotes significance at the 0.01, 0.05, 0.10 level.*

2.6 Summary and Discussion

This paper provides a quantitative review of the literature examining the impact of idiosyncratic information on firms' expected rate of returns. For the links between *Precision*/*Quantity* and CoE meta-analytic results are provided and for the link with *Asymmetry* descriptive statistics only due to data constraints. In total, results of 113 unique papers are reconsidered and findings suggest that the association between firm-specific information and CoE is subject to moderate effects. This is indicated by insignificant average effect sizes (*Precision*: -0.048 ± 0.095; *Quantity*: -0.066 ± 0.124) and a notable amount of studies (29%) which only find conditional results for the link between *Asymmetry* and CoE.

The empirical measurement of both CoE and information attributes partially explain these mixed results. First, the conjectured positive relationship between *Precision* and CoE is only significant in studies using *non*-accrual quality proxies for *Precision* and RFB/VMB proxies for CoE ($\bar{r}$: -0.063 ± 0.020); that is, the link between information precision and expected returns seems trivial in asset pricing tests (REAL proxies) or when accrual quality metrics are used. Second, almost all VMB studies confirm the positive association between *Asymmetry* and CoE (86%), but there is notable variation in the conclusions reached by REAL studies (reject: 14%; mixed: 38%; confirm: 48%). Third, mixed results for the link between *Quantity* and CoE are moderated by disclosure type (financial vs. partial-/non-financial) and disclosure settings insofar as firms' in comparatively weakly regulated countries tend to enjoy between 2 to 4 times greater CoE benefits from more expansive disclosure—depending on the type of disclosure—than firms in strongly regulated markets (such as the US or the UK).

Putting results into greater perspective, the following observations are particularly noteworthy. First, valuation model-based CoE estimates dominate the literature overall (i.e., 60 percent of all 138 observations use such proxies), but depending on which information link being examined the popularity of REAL, RFB and VMB proxies varies: *Asymmetry* studies mainly conduct asset pricing test or use realised returns to proxy for expected returns (75%), *Quantity* studies predominantly use VMB proxies (81%), with the greatest balance being observed for *Precision* studies (REAL: 33%; VMB: 54%). Interestingly, traditional RFB estimates are rarely used to forecast CoE (*Precision*: 13%; *Asymmetry*: 0%; *Quantity*: 10%), but of course are implicitly relied on in asset pricing tests. Second, there is large variety of different proxies for the information attributes, and results crucially depend on which proxy is used. For instance, the debate on the impact of *Precision* on firms' CoE is mainly an argument about the market pricing of accrual quality; the association between *Asymmetry* and CoE stems from the controversy over the pricing of PIN; and the impact of *Quantity* on CoE varies by disclosure types examined.

Given these results, my main two conclusions are as follows. First, wide variation in the empirical measurement of CoE and information attributes across studies hinders the assessment of the relative importance of *Precision*, *Asymmetry* and *Quantity* as determinates of firms' expected rate of returns; that is, which of the three attributes has comparatively greater CoE relevance cannot be answered conclusively by this study. Therefore, a comprehensive research design that allows to concurrently examine the different information attributes and CoE measures within one empirical model seems required to disentangle the underlying complexity between idiosyncratic information and expected rate of returns. My paper on "The Impact of Idiosyncratic Information on Expected Rate of Returns: A Structural Equation Modelling Approach" (Section 3) offers such a methodology and contributes new evidence to the debate on the pricing of information risk.

Second, VMB measures are highly popular and frequently used in the literature; hence, extant findings crucially hinge upon the empirical soundness of those measures. While convincing evidence exists that VMB proxies are indeed better measures of CoE than RFB and REAL (Botosan and Plumlee (2005), Botosan et al. (2011), Lee et al. (2010, 2015)), one should not neglect concerns raised in regard to the construct validity of these proxies. On the one hand, analyst-based ICC estimates tend to be upward biased (Easton and Monahan, 2005, 2016), which may distort findings (Hwang et al., 2013, p. 165); on the other hand, ICC estimates are mostly unavailable for young, small and financial distressed firms—the sort of firms which would be of "greatest interest to researchers examining issues related to information asymmetry, earnings quality, and disclosure where an ICC approach is used most often" (Li and Mohanram, 2014, p. 1153). In seminal work, Hou et al. (2012) and Li and Mohanram (2014) recommend the use of mechanical earnings forecasts to overcome these analyst-based deficiencies and to increase applicability of the ICC methodology; however, the extent to which their models are valid for the smallest, youngest and least followed firms in capital markets has not yet been examined. My paper on "Implied Cost of Capital and Cross-Sectional Earnings Forecasting Models: Evidence from Newly Listed Firms" (Section 4) fills this gap.

2.7 Appendix

Appendix 2.1: Journal Index

Journal	Abbreviation	ABS 2015
Academy of Taiwan Business Management Review	ATBMR	n/a
Accounting in Europe	AIE	2
Accounting Review	**TAR**	**4**
Accounting, Organizations and Society	**AOS**	**4**
Advanced Science Letters	ASL	n/a
Advances in Accounting, incorporating Advances in International Accounting	AIA	2
Applied Economics	AE	2
Asian Journal of Business and Accounting	AJBA	n/a
Asian Review of Accounting	ARA	2
Asia-Pacific Journal of Accounting & Economics	APJAE	2
Asia-Pacific Journal of Financial Studies	APJFS	n/a
Australian Accounting Review	AAR	2
Australian Journal of Management	AJM	n/a
Business Research	BR	n/a
China Finance Review International	CFRI	1
Contemporary Accounting Research	**CAR**	**4**
Corporate Social Responsibility and Environmental Management	CSREM	1
Emerging Markets Finance and Trade	EMFT	n/a
European Accounting Review	**EAR**	**3**
European Management Journal	EMJ	2
Global Journal of Business Research	GJBR	n/a
Industrial Management and Data Systems	IMDS	2
International Business Review	**IBR**	**3**
International Journal of Accounting	**IJA**	**3**
International Journal of Forecasting	**IJF**	**3**
International Review of Economics and Finance	IREF	2
International Review of Financial Analysis	**IRFA**	**3**
Journal of Accounting and Economics	**JAE**	**4**

Notes: Table continued next page.

Appendix 2.1: Journal Index (cont.)

Journal	Abbreviation	ABS 2015
Journal of Accounting and Public Policy	**JAPP**	**3**
Journal of Accounting Research	**JAR**	**4**
Journal of Accounting, Auditing and Finance	**JAAF**	**3**
Journal of Applied Accounting Research	JAAR	2
Journal of Banking and Finance	**JBF**	**3**
Journal of Business	JB	n/a
Journal of Business Finance and Accounting	**JBFA**	**3**
Journal of Business, Economics and Finance	JBEF	n/a
Journal of Corporate Finance	**JCF**	**4**
Journal of Economics, Finance and Administrative Science	JEFA	n/a
Journal of Empirical Finance	**JEF**	**3**
Journal of Finance	**JF**	**4**
Journal of Financial and Quantitative Analysis	**JFQA**	**4**
Journal of Financial Economics	**JFE**	**4**
Journal of Intellectual Capital	JIC	2
Journal of International Financial Markets, Institutions and Money	**JIFMIM**	**3**
Journal of Multinational Financial Management	JMFM	2
Journal of Risk and Insurance	**JRI**	**3**
Management Decision	MD	2
Management Science	**MS**	**4**
Management Science and Engineering	MSE	n/a
Managerial and Decision Economics	MDE	2
Managerial Finance	MF	1
North American Journal of Economics and Finance	NAJEF	2
Pacific Accounting Review	PAR	1
Quarterly Journal of Finance	QJF	1
Review of Accounting and Finance	RAF	2
Review of Accounting Studies	**RAST**	**4**
Review of Finance (European Finance Review)	**RF**	**4**
Review of Quantitative Finance and Accounting	**RQFA**	**3**
SSRN	SSRN	n/a

Notes: Journals in bold are categorised as high-tier journals in this study.

3 The Impact of Idiosyncratic Information on Expected Rate of Returns: A Structural Equation Modelling Approach

Abstract

This study applies an innovative structural equation modelling approach to examine the repercussions of idiosyncratic information on expected rate of returns. Using nine different proxies for cost of equity (*CoE*) and three different indicators for each information attribute (viz. *Quantity*, *Asymmetry* and *Precision*), this paper provides for a sample of 7,091 firms confirmatory evidence that companies with high (low) quality information environments enjoy relatively lower (higher) CoE than otherwise identical firms, with *Precision* and *Asymmetry* being of equal CoE relevance, while *Quantity* effects being economically negligible; however, findings also show that the significance of this impact decreases with firm size, maturity and profitability as well as market competition. Furthermore, informational differences between companies explain substantial variation in analyst-based implied cost of capital (ICC) estimates, but none in traditional risk factor-based return proxies (RFB), indicating that the former impound much more firm-specific information than the latter. Given the generally higher construct validity of ICC over RFB proxies, this suggests that incorporation of idiosyncratic information in the measurement of risk factor-based proxies might help improve the empirical soundness of those estimates.

3.1 Introduction

Extensive literature in accounting and finance investigates the extent to which idiosyncratic information affects price formation and return structures in capital markets. This line of inquiry commonly examines the proposition that firms with high (low) quality information environments should enjoy relatively low (high) costs of equity (e.g., Easley et al. (2002); Francis et al. (2005a)). Analytical work by Easley and O'Hara (2004, hereafter: EO) and Lambert, Leuz and Verrecchia (2012, hereafter: LLV) models this link elegantly and nominates a set of information attributes that characterise the overall quality of a firm's information environment (viz. *Quantity*, *Precision* and *Asymmetry*); however, the empirical validity of these models is subject to great debate.

Given on the one hand that the construct validity of different cost of equity (CoE) proxies is an ongoing debate in itself (e.g., Botosan and Plumlee (2005), Easton and Monahan (2016)), and on the other hand that proxies for the information attributes are large in numbers as informed by both accounting and finance research, the empirical literature is voluminous, and the conclusions reached vary widely depending on the proxies researchers use. Moreover, extant work examining the association between firm-specific information and expected rate of returns neglects the fact that interrelations between *Quantity*, *Precision* and *Asymmetry* exist and that direct and indirect paths between the attributes and CoE are of varying importance (Figure 3.1 depicts the prevailing methodological approach in current studies).

With that in mind, the main objective of this paper is to consolidate and reconcile previous empirical findings by means of a structural equation modelling (SEM) approach. More specifically, the first part of this paper analyses to what extent differences in firms' information environments impact investors' CoE expectations: holding the measurement of CoE constant while varying the informational part of the model. The second part examines the explanatory power of idiosyncratic information for different CoE measures: holding the informational part constant while varying CoE measurement.

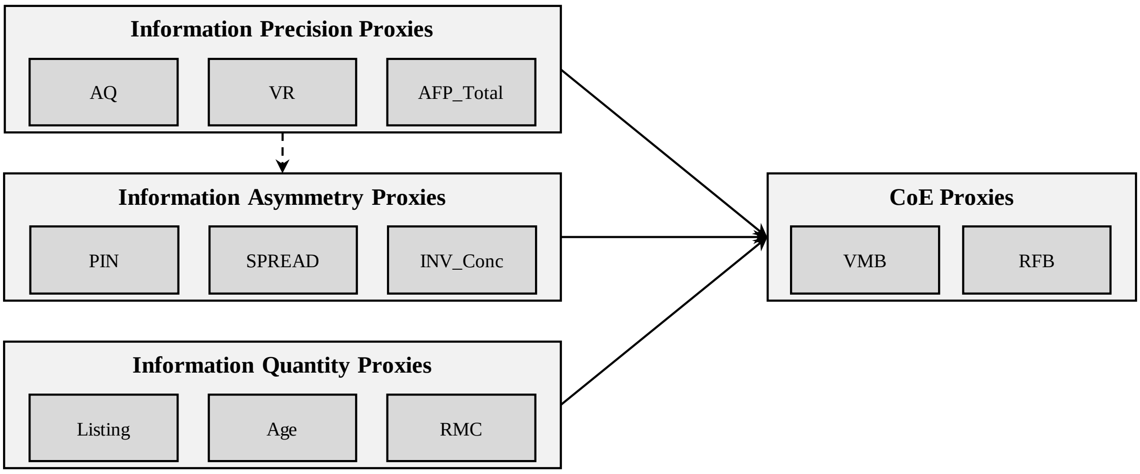


Figure 3.1: Prevailing Methodological Approach

Illustration of the methodological approach taken in extant work. Information attributes (i.e., Quantity, Precision and Asymmetry) are treated as directly measurable by means of different proxies and their impact on CoE is assessed by multivariate regression analyses. Dashed line between information precision and asymmetry indicate a recent study by Bhattacharya et al. (2012) who use path analysis to decompose the association between information precision and CoE into a direct and indirect path mediated by information asymmetry; however, path analysis maintains the assumption of direct, error-free measurability of the attributes by a single-indicator (one observed measure per attribute). Variable names as defined in Table 3.1.

There are two major advantages to this study's SEM approach vis-à-vis extant research designs in examining the CoE effects of firms' information environments. First, SEM allows measuring each information attribute by multiple proxies (i.e., factor analyses identify shared variability between the proxies as a measure of the information attributes), while previous studies only use one proxy per attribute (i.e., direct measurability of the information attributes is assumed); similarly, I estimate firms' expected rate of returns (*CoE*) from both risk factor-based (RFB) and valuation model-based (VMB) CoE proxies. As a result, widely used proxies for *Quantity*, *Precision*, *Asymmetry* and *CoE* in previous studies are now concurrently accommodated and directly tested within one empirical model. Second, interrelations among information attributes and associations between the attributes and CoE are estimated simultaneously, which provides test results for the model as whole rather than just for path coefficients individually as in previous work (Figure 3.2 shows the conceptual model of this paper).[23]

[23] Appendix 3.1 provides a more detailed description of the SEM methodology and why it is an appropriate approach in studying the link between idiosyncratic information and cost of equity.

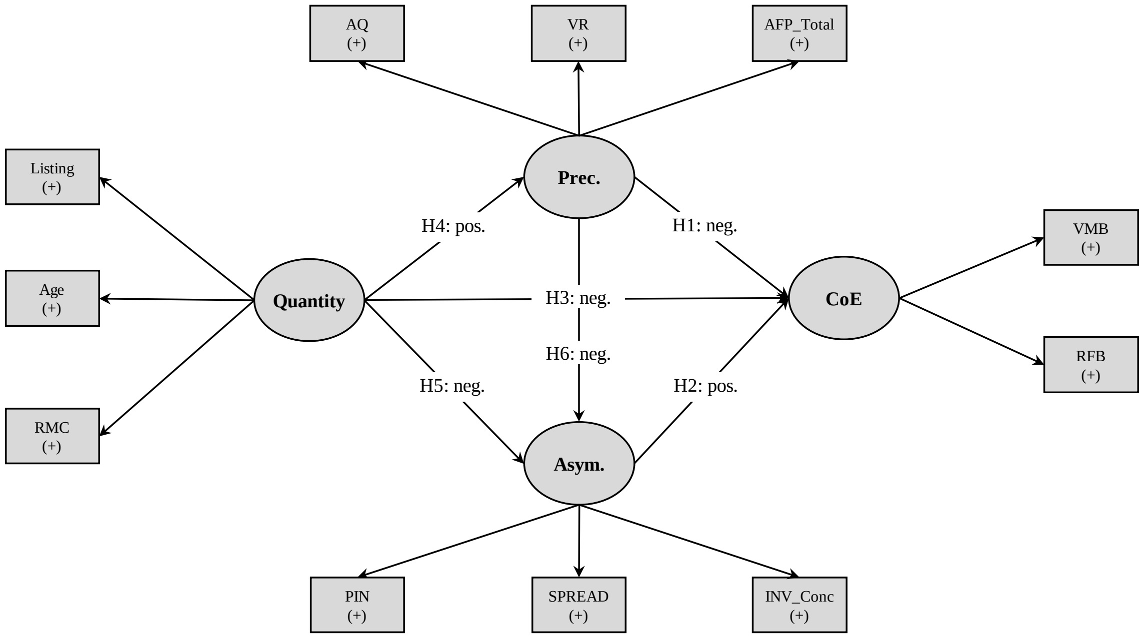


Figure 3.2: Conceptual Model

Oval figures indicate latent constructs (i.e., information attributes) which are only indirectly measurable through their impact on observable indicator variables/proxies (rectangular figures). One-headed arrows indicate regression relationships, and the end of each arrow indicates the dependent construct. Plus (minus) signs represent positive (negative) predicted relations between information attributes and indicator variables. Pos.(Neg.) indicates hypothesised association between latent constructs. Error terms are suppressed. Variable names as defined in Table 3.1.

Overall, this study makes three contributions to the literature. First, using a SEM approach extends Bhattacharya et al. (2012, hereafter: BEOS)—who use path analysis—from a methodological perspective; that is, SEM acknowledges that information attributes are latent constructs which are only indirectly measurable by two or more proxies, while path analysis maintains the less realistic assumption of direct, error-free measurability of the respective attributes via a single empirical measure. Having this conceptual benefit, I use three different proxies (*alias* indicators) for each information attribute to infer firms' *Quantity*, *Asymmetry* and *Precision*, and test nine different CoE measures.[24] Hence, only the "good" part of each proxy is retained to contribute to the measurement of its assigned attribute, which mitigates bias in estimated structure coefficients (emanating from the errors-in-variable problem) and increases the empirical robustness of my findings (Rao, 1973).

[24] I test three valuation model-based (i.e., implied cost of capital) estimates, three traditional risk factor-based ones (e.g., rFF3), and three market sentiment (i.e., VIX) augmented proxies.

Second, results from the first part of my analysis bring new evidence to the debate over the pricing of information risk (e.g., Core et al. (2008), Mohanram and Rajgopal (2009), Shevlin (2013)). Findings corroborate the proposition that firms with higher (lower) quality information environments enjoy relatively lower (higher) cost of equity than otherwise identical companies, with *Precision* (-0.447) and *Asymmetry* (0.437) being of equal CoE relevance, while *Quantity* (0.139) effects are economically negligible (structure coefficients in parentheses). However, the strength of this impact varies with firm size, maturity and profitability as well as market competition about companies' shares (i.e., stock liquidity). This can be inferred from decreasing levels of explanatory power as one moves from AMEX (CoE R^2: 0.603) to NASDAQ (0.339) to NYSE (0.182) stocks, with the latter stocks being much larger, older and more profitable than the former two. In addition, the direct association between *Asymmetry* and *CoE* for high competition stocks is almost four times weaker (0.058) than for low competition stocks (0.209) and about 1.5 times weaker (high: -0.519; low: -0.328) for *Precision*; this can be seen as an empirical extension of the LLV (2012) model insofar as it is not only *Asymmetry*, but also *Precision* effects which tend to decrease as market competition, and thus, stock liquidity, increases.

Third, findings from the second part of my analysis contribute to the discussion about the empirical validity of different CoE measures (e.g., Botosan and Plumlee (2005), Easton and Monahan (2016)). The SEM model explains a significant amount of variation in VMB (i.e., implied cost of capital) proxies (R^2: 0.244), but none of the variation in traditional RFB measures (i.e., *rCAPM*, *rFF3*, *rFF4*). As VMB proxies derive from analyst forecasts (hence, impounding more idiosyncratic information than RFB measures), it is unsurprising that *Quantity*, *Asymmetry* and *Precision* explain variation in the former much better than in the latter. However, combining this observation with the fact that VMB generally show greater construct validity than RFB proxies (Botosan et al. (2011), Lee et al. (2010)), implies that firm-specific information might be the missing "explanatory link" between the two. Differently stated, the incorporation of idiosyncratic information in the measurement of risk factor-based estimates might improve those proxies' construct validity and thus help reconcile prevailing performance differences with implied cost of capital (ICC) estimates.

To the best of my knowledge, this is the first study which provides simultaneous evidence on how *Quantity*, *Precision* and *Asymmetry* affect firms' *CoE* within a comprehensive empirical model and, as such, might be acknowledged as a "first step" (Beyer et al., 2010, p. 309) in the right direction to disentangle the underlying complexity between firm-specific information and expected rate of returns.[25]

The remainder of this paper proceeds as follows. Section 3.2 provides a theoretical framework for the conceptual model and derives the research hypotheses. Section 3.3 describes the measurement of the indicators and cost of equity proxies along with a description of the estimation and assessment of the study's SEM models. The main analyses are conducted in Section 3.4, and concluding remarks are provided in Section 3.5.

3.2 Theoretical Background and Research Hypotheses

3.2.1 Market Microstructure and Information-Based Models

The idea behind the conceptual model in this study (see Figure 3.2) is widely consistent with seminal work on the interrelation between idiosyncratic information and firms' CoE by Easley and O'Hara (2004), who are the first scholars to present an information-based model that bridges the gap between market microstructure and asset pricing theory.[26] While it should not be overlooked that incomplete information models (Merton, 1987),

[25] In a recent working paper, Hinson and Utke (2016) give an introduction on the applicability of SEM in archival accounting research and provide some empirical results for the impact of disclosure on cost of capital. Their paper also builds on BEOS (2012); however, my paper differs from theirs in at least four accounts; (1) their paper has a predominate focus on the impact of disclosure on CoE and as such tests only two hypotheses. In contrast, my paper examines six hypotheses which pertain to the relations between *Quantity*, *Precision*, *Asymmetry* and *CoE*; (2) their analysis is limited to one SEM model, while I provide six different SEM specifications with a particular focus on differences across groups of firms; (3) this study offers insights on the relative importance of idiosyncratic information as an explanatory variable for different CoE measures—which theirs do not; and (4) my paper operates a much larger sample (60,995 vs. 7,642 observations) and spans a much longer sampling period (1993-2010 vs. 2000-2010) than their paper.

[26] Most past economic research (including the asset pricing literature) is mainly focused on the study of price equilibria, with only little – if any – consideration given on how these equilibria are actually attained. The rather new field of market microstructure research attempts to overcome this negligence in that the decomposition of the price building process into its underlying determinants is at the forefront of this line of inquiry. One of the more recent approaches in this attempt is to study the "information-aggregation properties of prices and markets" by means of information-based models. A central objective of this particular subfield of market microstructure research is to answer how different sets of information may impact the price formation and, thus, the return structure in capital markets (O'Hara, 1995).

asymmetric information models (Admati, 1985) and liquidity effect models along the work of Amihud and Mendelson (1986), Brennan et al. (1998) and Diamond and Verrecchia (1991) propose insightful theoretical explanations for the impact of information on the price and return behaviour of assets, it is EO (2004) who model the "effects of estimation risk and information asymmetry within a unified framework" (Botosan et al., 2004, p. 239).

3.2.2 Impact of Information Attributes on Cost of Equity

EO (2004) demonstrate that a firm's CoE decreases in *Precision* (i.e., decreases in accuracy of the available information about the future value of the firm) and increases in *Asymmetry* (i.e., increases in the fraction of uninformed investors and private signals about the future value of the firm). The economic intuition behind their model is as follows. First, investors demand to be rewarded for bearing uncertainty about a firm's future prospects, which stems from imprecise information given to them, implying that firms that disclose higher quality financial and non-financial data to investors can benefit from reduced CoE. Second, uninformed investors (who only have access to public information) require compensation for "losing out" against privately informed investors (who have access to both public and private information) when making investment decisions. The higher this informational disparity between these two groups is, the larger the CoE premium induced by *Asymmetry*; however, LLV (2012) illustrate that in perfectly liquid markets in which both the informed and uninformed act as price takers, asymmetric information has no effect on the CoE over and above its impact on average precision about the future value of the firm. This subtle extension of the EO (2004) model entails that the degree of market competition represents a crucial conditioning variable in empirical settings. [27]

By reviewing some notable empirical studies that draw upon the EO (2004) and/or LLV (2012) model, I validate the hypothesised links between the information attributes and CoE in my conceptual model, reveal the most commonly used proxies for the respec-

[27] A detailed description of the EO (2004) and LLV (2012) model is provided in Appendix 3.2 and Appendix 3.3, respectively.

tive information attributes in extant work (which are then used as measures for the information attributes in later analyses) and present evidence about the extent to which each proxy is associated with CoE on a stand-alone basis. The empirical measures for the information attributes are introduced and discussed below, and the procedure for estimating them is in the methodology section.

3.2.2.1 Information Precision and Cost of Equity

The theoretical link between *Precision* and *CoE* has received much attention in empirical studies, resulting in two important strands of research. The first strand uses earnings quality metrics, and the second one utilises security analysts' forecasts as empirical measures of information precision.

Earnings Quality. Earnings quality metrics are a natural choice for *Precision* indicators, given that "higher quality earnings provide more information about the features of a firm's financial performance" than lower quality earnings (Dechow et al., 2010, p. 344). Extensive empirical evidence shows that investors regard earnings as an important source of information about the performance of a company which substantiates their inclusion as measures of information precision (e.g., Biddle et al. (1995), Charitou et al. (2001), Francis et al. (2003), Liu et al. (2002), Strong (1993), Strong and Walker (1993)). In a seminal study, Francis et al. (2004, hereafter: FLOS) find that for seven different earnings metrics, higher earnings quality leads to lower CoE. The strongest effects are observed for the *accrual quality (AQ)* and *value relevance (VR)* metrics, which I use as an accounting-based and market-based indicator for earnings quality, respectively. Recent studies corroborate the empirical validity of *AQ* and *VR* and demonstrate that higher earnings quality, and thus lower earnings management, lead to favourable cost of equity effects (e.g., Aboody et al. (2005), Barth et al. (2013), Francis et al. (2005a, hereafter: FLOS), Gray et al. (2009), Kim and Qi (2010), Ogneva (2012)).[28]

[28] Although Core et al. (2008) strongly question the validity of this evidence, the authors acknowledge that accruals quality is certainly not irrelevant as they find a positive association between AQ and market beta. This result is indicative of accrual quality having (at least) an indirect effect on CoE through market beta. Ogneva (2012) further reconciles these opposing views in that she relies on Campbell and Shiller's (1988a, 1988b) return decomposition to show that accruals quality is positively associated with future cash flow shocks and as such hinders the detection of an association between AQ and realised returns. What is more, Gray et al. (2009) replicate the methodology of both Francis et al. (2005) and

Security Analyst Forecasts. Barron et al. (2005) demonstrates that the precision of analyst forecasts is a reliable measure for the general information precision of sophisticated investors; hence, several papers suggest the use of analyst forecast-based proxies as measures of *Precision* (e.g., Barry and Brown (1985), Barron and Stuerke (1998)). Drawing upon the framework in Barron et al. (1998), which shows that observable characteristics of analyst forecasts allow for inferences about the public, private and total precision of analysts' information sets, we employ *total analyst forecast precision* (*AFP_Total*) as an indicator for information precision in this study. This choice is motivated by robust evidence for a negative association between total analyst forecast precision and CoE (Barron et al. (2012); Botosan and Plumlee (2013)). Given this empirical evidence along with the analytical insights from the EO (2004) model, I formulate Hypothesis 1.

H1: The higher (lower) the information precision of a firm, the lower (higher) its CoE.

3.2.2.2 Information Asymmetry and Cost of Equity

Numerous studies evaluate the proposition that an increase in informational disadvantages between groups of investors provokes unfavourable CoE effects. This literature can be categorised into two different streams. The first derives its *Asymmetry* proxies from market microstructure data, and the second—which is based on the LLV (2012) prediction that the impact of information asymmetry on CoE is conditional on the degree of market competition—puts forward firm-ownership-based measures.

Market Microstructure. Bid-ask spreads and PIN scores are two prevalent measures of *Asymmetry*. The former is an indirect proxy for information asymmetry (i.e., bid-ask spreads increase as market-makers are exposed to greater adverse selection problems), with ample empirical evidence documenting a positive relationship between spread-based proxies and firms' CoE (e.g., Amihud and Mendelson (1986), Bhattacharya et al. (2012), Levi and Zhang (2015)). In contrast, the latter is a direct measure of information asymmetry in that it uses order-imbalances to infer the probability that the next trade order is from a privately informed investor, i.e., based on private information. Seminal evidence

Core et al. and find for a sample of Australian firms that, irrespective of the asset pricing approach taken, AQ is priced in security markets.

establishing a positive association between PIN scores and expected rate of returns is contributed by Easley, Hvidkjaer and O'Hara (2002) and verified in—among others—Brennan et al. (2016), Duarte et al. (2008), Easley et al. (2010) and Yan and Zhang (2014).[29] Given this sheer volume of positive empirical findings, I use both *PIN* scores and bid-ask spreads (*SPREAD*) as measures of information asymmetry in this study.

Market Competition. Following the LLV (2012) model, the impact of information asymmetry on CoE is conjectured to be conditional on the degree of market competition. Three recent empirical studies scrutinise this new analytical insight, but they deliver mixed empirical results. While Armstrong et al. (2011) and Akins et al. (2012) provide evidence supporting the prediction of a diminishing impact of information asymmetry on the CoE as markets approach perfect competition, Barron et al. (2012) show that, irrespective of market setting, information asymmetry has a significant unfavourable effect on firms' expected rate of returns. Reconciliation of these opposing results is hindered by the variety of proxies employed in each paper.[30] However, consistent with the cited studies, I hypothesise that the degree of market competition is indicative of the degree of information asymmetry, i.e., greater competition among informed investors leads to a quicker revelation of private information in prices and consequently reduces informational disadvantages between investors.[31] Because the dispersion of shares among investors tends to be the main determinant of market competition (Armstrong et al., 2011, p. 11), I use the investor concentration measure (*INV_Conc*) first suggested by Akins et al. as an indicator for market competition, with greater concentration being conjectured to be associated

---

29 Mohanram and Rajgopal (2009) conclude that "there is not much evidence to support the interpretation that information risk, proxied by PIN, is a source of priced information risk." However, they acknowledge that while their paper suggests "PIN is not priced risk, it is difficult to make more general statements about the pricing of information risk since information risk can […] be proxied by different empirical variables" (ibid., p. 241). By using a spread-based proxy as a complementary indicator, I address their concern with respect to PIN scores.

30 First, Armstrong et al. and Barron et al. share a common proxy for market competition (number of total shareholders), but use distinct proxies for information asymmetry; i.e., Armstrong et al. deploy a market microstructure proxy (bid-ask spreads) while Barron et al. apply an analyst forecast-based proxy. Second, Akins et al. use number of institutional investors as well as an investor concentration measure to proxy for market competition; and proxy for information asymmetry by means of the information and non-information asymmetry component of bid/ask spreads as well as adjusted PIN scores.

31 For instance, Armstrong et al. (2011, p. 9) argue that "if the number of shareholders increases, the slope of the price curve caused by any information asymmetry decreases...." See also Foster and Viswanathan (1994).

with *less* competition and thus *greater* information asymmetry. Taken together, the empirical findings support Hypothesis 2.

H2: The higher (lower) the information asymmetry between investors, the higher (lower) the firm's CoE.

3.2.2.3 Information Quantity and Cost of Equity

The EO (2004) and LLV (2012) model primarily examine the CoE effects of *Precision* and *Asymmetry*.[32] However, I also include *Quantity* as an information attribute in my model because the estimation risk literature shows that if the amount of information about a firm is low, investors have difficulties accurately estimating the return parameters of this particular firm, which makes it a riskier investment and hence induces higher CoE (e.g., Clarkson et al. (1996), Kumar et al. (2008), Lewellen and Shanken (2002), Zhang (2006)).

Supported by evidence in (Barry and Brown, 1984, 1985)—who show that the period of listing is negatively associated with risk-adjusted returns—and Clarkson and Thompson (1990)—who find that average beta risk for newly issued firms declines over the next few months of listing—I use *Listing* as an indicator for *Quantity*. Given that firms' operating histories tend to be positively associated with the amount of information prior to their listing (Clarkson and Satterly (1997), Ecker (2014), Lee et al. (2003)), I consider firm age (*Age*) as a complementary indicator.

A large stream of research investigates if there exists a negative association between corporate disclosure and expected rate of returns.[33] Disclosure scores are intuitive proxies for information quantity; however, I refrain from using them given that they tend to be noisy indicators of *Quantity* (i.e., disclosure scores tend to capture both a quantity and

---

32 Only an exogenous increase in prior information and/or private and public information in the EO and LLV model lead to a decrease in the CoE, *ceteris paribus*.

33 See, amongst others, Campbell et al. (2014), Core et al. (2015) for an analysis of mandatory disclosure; Botosan and Plumlee (2002), Francis et al. (2008) for voluntary disclosure; Baginski and Rakow (2012), Evans (2016) for financial disclosure; and Dhaliwal et al. (2011), Ng and Rezaee (2012) for non-financial disclosure.

quality dimension) and there is no consensus on their measurement.[34,35] Instead, I apply a self-constructed relative media coverage (*RMC*) index based on prior evidence which shows that information quantity increases with media coverage (Kross and Schroeder, 1989). I summarise the above in Hypothesis 3.

H3: The larger (smaller) the quantity of available information about a firm, the lower (higher) its CoE.

3.2.3 Interrelations between Information Attributes

The opportunity to model interrelations between latent constructs and treat them simultaneously as dependent and independent variables is a key benefit of SEM models. This facilitates the objective of this study to consolidate the extensive empirical evidence in extant research within one model. As shown in Figure 3.2, I hypothesise that beyond the direct paths, indirect paths from the attributes to the CoE exist. While the EO (2004) and LLV (2012) models are silent on these indirect associations, empirical findings on the relation among the information attributes warrant their inclusion.

3.2.3.1 Quantity and Precision

Sufficient evidence supports the conjecture that information quantity is positively associated with information precision and, furthermore, that causality runs from *Quantity* to *Precision*. For instance, disclosure levels tend to be positively associated with analysts' forecast precision (e.g., Lang and Lundholm (1996) demonstrate that increased disclosure levels lead to more accurate and less dispersed forecasts; Byard and Shaw (2003) report that higher quality disclosure increases both private and public analysts' forecast precision) which supports the notion of more disclosure (i.e., more quantity) leading to greater accuracy in analysts' forecasts (i.e., higher information precision).

---

34 For instance, Cheng et al. (2006, p. 179) state that "while prior empirical research has used the *quantity* of disclosure as a proxy for the *quality* of disclosure quality, in many cases disclosure quantity and quality are not separable information attributes."

35 Some authors use self-constructed disclosure scores (e.g., Botosan (1997), Kothari et al. (2009)), while others rely on commercially available ones (e.g., Healy et al. (1999), Richardson and Welker (2001)), and yet others use simple dummy variables to distinguish between disclosing and non-disclosing firms (e.g., Ogneva et al. (2007), Cao et al. (2017)).

In addition, disclosure levels show a significant association with earnings quality, although the sign of the association is debated. Some studies find a positive (complementary) relation, i.e., firms with better earnings quality disclose more expansively (e.g., Francis et al. (2008), Waymire (1985), Cox (1985), Imhoff (1978)), while others find a negative (substitutive) relation (e.g., Lang and Lundholm (1993), Tasker (1998)). These contradictory results might stem from the fact that proxies for disclosure levels tend to capture both a quantity and quality dimension.[36] However, irrespective of the sign of the association, the key insight is that disclosure levels tend to have a significant effect on firms' earnings quality, which in turn supports the conclusion of *Quantity* influencing *Precision*. Therefore, I formulate Hypothesis 4.

H4: The higher (lower) the quantity of available information about a firm, the higher (lower) the firm's information precision.

3.2.3.2 Quantity and Asymmetry

Empirical evidence conclusively shows that as the quantity of information about a firm increases (proxied by disclosure levels in the following studies), the fraction of privately informed investors (information asymmetry) decreases. This holds true irrespective of whether spread-based proxies (Healy et al. (1999), Heflin et al. (2005), Welker (1995)) or PIN-based proxies (Brown and Hillegeist (2007), Brown et al. (2004)) as measures of information asymmetry are applied. These results also advocate the assumed causality from *Quantity* to *Asymmetry*. Derived from these findings, I express Hypothesis 5.

H5: The higher (lower) the quantity of available information about a firm, the lower (higher) the information asymmetry between investors.

3.2.3.3 Precision and Asymmetry

Recent findings lend support to the notion of information asymmetry being negatively associated with information precision, where causality is assumed from the former to the

[36] Say, one proxy for disclosure level mainly captures the quality dimension of disclosure, while another one has a greater exposure to the quantity dimension. Then the former is by construction more likely to be positively associated with earnings quality (given that it measures the same underlying concept) than the latter proxy which mainly captures the quantity dimension of disclosure.

latter information attribute. For instance, Bhattacharya et al. (2013) find that earnings quality (measured by the *AQ* metric) is negatively associated with the adverse selection component of bid-ask spreads, and Bhattacharya et al. (2012) conclude that "there appears to be a limited (in magnitude) feedback path from information asymmetry [proxied as the adverse selection component of bid-ask spreads] to earnings quality…." (p. 472). This lends support to Hypothesis 6.

H6: The higher (lower) the information precision of a firm, the lower (higher) the information asymmetry between investors.

In summary, this study's conceptual model is informed by both contemporary analytical and empirical evidence. The direct relations between the attributes and CoE follow from the theoretical models developed by EO (2004) and LLV (2012), while the interrelations among the attributes, and therefore, the indirect paths with CoE, are motivated by recent findings in empirical work. Ideally, the comprehensiveness of the hypothesised model not only provides new insights on the extent to which (and under what circumstances) differences in firms' information environments affect investors' return expectations, but also on the relative importance of firm-specific information as an explanatory variable for different CoE measures.

3.3 Methodology

3.3.1 Empirical Measures of Indicators and Cost of Equity

Table 3.1 provides an overview of the constructs and their respective indicators applied in my model and the databases used to estimate them. To conserve space, detailed descriptions of how each indicator is estimated are relegated to Appendix 3.4 for *Quantity*, Appendix 3.5 for *Precision*, Appendix 3.6 for *Asymmetry and* Appendix 3.7 for *CoE* measures. All variables which are not derived from yearly data (such as *Listing, Age, RMC, AQ, VR*), but estimated from quarterly or daily information instead (such as *AFP_Total, PIN, SPREAD, INV_Conc*) are time-series averages of lead, lag and current observations centred around the fiscal year-end of each firm. This mutes potential noise in the proxies surrounding the release of fiscal year-end information and is consistent with empirical work relying on comparable data (e.g., Bhattacharya et al. (2012), Botosan and Plumlee (2013)). Similarly, risk factor-based proxies (e.g., rFF4) are calculated six months after firms' fiscal year-ends, and valuation model-based proxies (e.g., rPEG) are derived from the first available consensus analyst forecasts immediately made after firms' earnings announcements.

Table 3.1: Overview of Constructs and Indicators

Construct Indicator	**Name**	**Association with construct**	**Database**
Quantity			
Period of listing (IPO)	Listing	pos.	SDC Platinum, Osiris & CRSP
Firm age (incorporation)	Age	pos.	Osiris
Relative media coverage	RMC	pos.	Factiva
Precision			
Accrual quality	AQ	pos.	Compustat
Earnings value relevance	VR	pos.	Compustat & CRSP
Total analyst forecast precision	AFP_Total	pos.	I/B/E/S
Asymmetry			
Probability of an informed trade	PIN	pos.	Stephen Brown
Bid/Ask spread	SPREAD	pos.	CRSP
Investor concentration	INV_Conc	pos.	CDA/Spectrum (s12)
CoE – Risk Factor-Based (RFB)			
Capital Asset Pricing Model	rCAPM	pos.	
Fama-French 3-factor model	rFF3	pos.	Kenneth French & CRSP
Carhart's 4-factor model	rFF4	pos.	
CoE – Risk Factor-Based + VIX (FVIX)			
Capital Asset Pricing Model + FVIX	rFVIX	pos.	
Fama-French 3-factor model + FVIX	rFVIX3	pos.	Kenneth French, Alexander Barinov & CRSP
Carhart's 4-factor model + FVIX	rFVIX4	pos.	
CoE – Valuation Model-Based (VMB)			
Price-Earnings-Ratio	rPE	pos.	
Price-Earnings-Growth	rPEG	pos.	I/B/E/S & Compustat
Abnormal Earnings Growth Model	rAEGM	pos.	
Future Realised Returns (FRR)			
Annualised 12-month Buy-Hold-Return	ret12	pos.	
Annualised 24-month Buy-Hold-Return	ret24	pos.	CRSP
Annualised 36-month Buy-Hold-Return	ret36	pos.	

Table shows indicators per information and CoE construct including variable names, predicted association with the respective construct and the databases used to measure the indicator.

3.3.1.1 Information Quantity

My three indicators to infer information quantity are period of listing (*Listing*), firm age (*Age*) and relative media coverage (*RMC*). Period of listing (firm age) is the number of years since a firm's initial public offering (incorporation) and is calculated as the difference between the year of fiscal year-end and the year of the IPO (incorporation). Inspired by the work of Beretta and Bozzolan (2004, 2008) and Beattie et al. (2002, 2004), the idea behind the RMC index is that variations in media coverage—which are unexplained by industry membership and firm size (i.e., residuals)—are a good proxy for firm prominence that, in turn, tends to be positively associated with higher information quantity (Kross and Schroeder, 1989); thus, the RMC index is positive (negative) for firms that enjoy more (less) media coverage than the expected average firm in the same industry and of similar size. All indicators are assumed to be positively associated with *Quantity*.

3.3.1.2 Information Precision

I measure information precision by two earnings quality indicators, *accrual quality* and *earnings value relevance*, and one indicator for the precision of analyst forecasts, *AFP_Total*. The accrual quality metric (*AQ*) is based on the McNichols (2002) modification of the Dechow-Dichev (2002) model; however, I deviate from previous studies and multiply the *AQ* metric by negative one to establish a more intuitive relation with *Precision*: the lower (higher) the information precision of a firm, the lower (higher) its accrual quality. The measurement of the earnings value relevance metric (*VR*) in this study is similar in estimation and interpretation to the *AQ* metric above: the lower (higher) the information precision of a firm, the less (more) value relevant its earnings.[37] That is, *VR* is assumed to be positively associated with *Precision*.

As with respect to the security analyst-based indicator (*AFP_Total*), I draw upon seminal work in Barron et al. (1998), that shows that observable characteristics of analysts' forecasts (viz. forecast dispersion, squared error in the mean forecast and the number of forecasts) can be used to make inferences about the degree of information precision of the public and private information sets available to security analysts. The sum of those two components reflects the total precision of analysts' information sets and substitutes well for the general information precision of sophisticated investors (Barron et al., 2005). The empirical estimation of *AFP_Total* in this study is widely consistent with Botosan and Plumlee (2013) and hypothesised to be positively related with *Precision*.

[37] Value relevance of earnings is expressed as the degree to which both a firm's earnings and change in earnings explain its stock returns, where greater explanatory power indicates more transparent and value relevant earnings, respectively (Barth et al. (2013), FLOS (2004)).

3.3.1.3 Information Asymmetry

I use bid-ask spreads (*SPREAD*) and probability of informed trading (*PIN*) scores as two microstructure proxies for information asymmetry. Copeland and Galai (1983) and Glosten and Milgrom (1985) formally show that bid-ask spreads are valid proxies for the exposure of market-makers to the adverse selection problem and as such capture well the degree of information asymmetry between informed and uninformed investors.[38] Consistent with Stoll (1978), I use the average of simple daily percentage spreads as an indicator for firms' bid-ask spreads (*SPREAD*). PIN scores are firm-specific proxies for information asymmetry that measure the probability that the next trade order is from a privately informed investor, where larger PIN scores signify larger information asymmetry (Easley et al. (1996, 1997), Brown and Hillegeist (2007)). The underlying notion of the PIN model is that while it is impossible to directly observe which trades are based on private information, one can use imbalances between buy and sell orders to infer the probability of information-based trading for a given stock. I obtain quarterly PIN scores from Stephen Brown's website.[39]

These two microstructure proxies are complemented by investor concentration (*INV_Conc*) as a measure of market competition. The estimation of *INV_Conc* is similar to Akins et al. (2012), but I use information on mutual fund holdings instead of institutional investor holdings due to data restrictions. It is conjectured that higher values of concentration denote less competition in the trading of firms' stocks; hence, *INV_Conc* is assumed to be positively associated with information asymmetry.[40]

[38] Jack Treynor, publishing under his pseudonym Walter Bagehot (1971), gives an intuitive explanation as to why greater bid-ask spreads are associated with greater information asymmetry. In essence, he argues that market makers are aware of the fact that they lose against the privately informed and, therefore, have to cover their loses by trading profitably with the uninformed. The only way they can avoid this adverse selection problem is to quote different prices for buying and selling orders, i.e. profiting from different bid and ask prices in order to maintain their own liquidity. This then implies that bid-ask spreads, which stem from adverse selection problems market makers are exposed to, are informative about the fraction of privately informed investors in the market.

[39] http://scholar.rhsmith.umd.edu/sbrown/pin-data

[40] Akins et al. multiply the index by negative one, so that higher values indicate *more* competition (p. 41); however, to maintain consistent interpretation between all of our indicators and asymmetry (viz. a conjectured positive association), we refrain from this modification.

3.3.1.4 Cost of Equity and Future Realised Returns

I calculate nine different proxies for firms' CoE: three risk factor-based (RFB) ones, three market sentiment (i.e., VIX) augmented proxies, and three valuation model-based (VMB) estimates (i.e., implied cost of capital). I use *rCAPM*, *rFF3*, and *rFF4* as my RFB proxies (Carhart (1997); Fama and French (1993); Lintner (1965); Mossin (1966); Sharpe (1964)) and re-estimate them with an additional risk-factor for expected market volatility (i.e., FVIX) to capture market sentiment: *rFVIX*, *rFVIX3* and *rFVIX4* (Ang et al., 2006).[41] Given that RFB proxies tend to be "imprecise estimates of the cost of equity" (Fama and French, 1997, p. 154), I follow previous work (e.g., Gordon and Gordon (1997); Botosan and Plumlee (2002); Claus and Thomas (2001); Gebhardt et al. (2001); Ohlson and Juettner-Nauroth (2005); Gode and Mohanram (2003)) and also estimate three implied cost of capital (ICC) estimates: *rPE, rPEG, rAEGM* (Easton, 2004).

In the first part of my analysis (Section 3.4.2 Results of SEM Analyses), the CoE construct is kept constant and inferred from one VMB (*rPEG*) and one RFB (*rFF4*) proxy. Those two proxies are chosen based on results from the second part of my analysis (Section 3.4.3 Results of CoE Analyses) where it is shown that *rPEG* and *rFF4* provide the most parsimonious and well-fitted SEM specifications among all tested CoE combinations (Figure 3.3 illustrates the interrelation between the two analytical steps). Moreover, inclusion of those two opposing CoE measures also results in a robuster estimate of expected returns in that CoE is now a measure of shared variance or commonality between the two (i.e., SEM controls for measurement error in the two proxies). While the RFB methodology assumes that past realised returns are a reliable guide for expected returns, VMB estimates derive from future dividend/earnings expectations and current market prices. Considering both proxies concurrently ensures that the final CoE construct captures—or at least controls for—both an *ex post* and an *ex ante* perspective.

I also calculate annualised 12, 24 and 36 months buy-and-hold returns, against which I validate the CoE measures. Each stock is assumed to have been bought six months after

[41] The notion underlying the FVIX factor is that companies with more (less) negative return sensitivity to VIX index changes have higher (lower) CoE. FVIX reflects the monthly excess return on a factor-mimicking portfolio that tracks daily changes in the VIX index (Barinov, 2013, p. 1880).

a firm's fiscal year-end and held for the subsequent 12, 24 and 36 months. These total returns are annualised (geometric mean) to make them comparable.

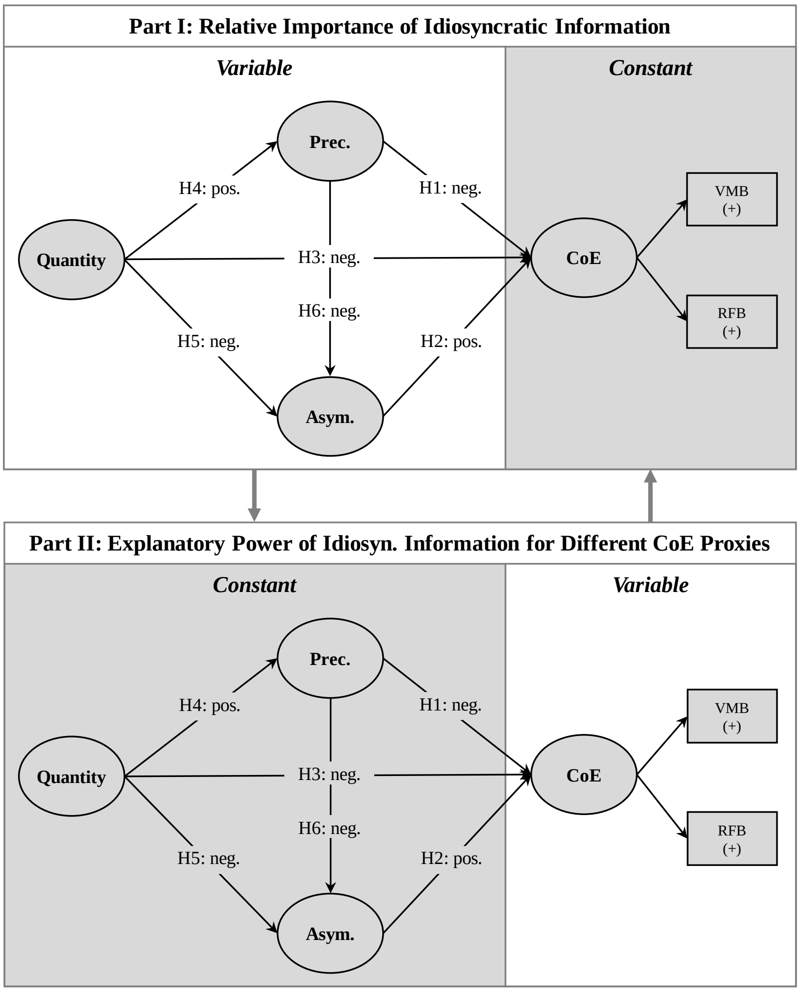


Figure 3.3: Feedback Loop between Analytical Steps

The figure illustrates the interrelation between the two analytical steps in this study. In the first step, the measurement of CoE is kept constant. In the second step, the measurement of the information part is kept constant. Results from the first step inform the constant part in the second step—and vice versa.

3.3.2 Sample and Data Selection

My sample includes all NYSE, AMEX and NASDAQ listed securities that are covered on the CRSP/Compustat Merged (CMM) database and have fiscal year ends from 1993-2010. This 18-year period is the longest possible sampling period for all necessary information. Data availability is greatest for RFB proxies; hence, I build our sample around all firms with available risk factor-based CoE estimates to maximises number of observations; this results in a final sample of 7,091 firms and 60,995 firm-years. Given that different methodologies and databases are required to estimate the respective indicators, sample size varies for each proxy. For instance, for only 1,887 of the 7,091 RFB sample firms *RMC* can be calculated (see Table 3.2). While varying sample size poses serious problems in OLS regressions, SEM copes reasonably well with missing data. Assuming that data are missing at random—a reasonable assumption for my sample—item non-response can be ignored, and maximum-likelihood missing data techniques ensure that parameter estimates are valid and efficient (Schafer and Graham, 2002).[42]

Overall, my sample is tilted towards larger and more profitable firms. The average sample return on assets (ROA) ranges from a low of 3.06 percent (for *SPREAD* and *RFB proxies*) to a high of 4.76 percent (for *Age*). Comparing these figures against the average median ROA of 1.78 percent for all CMM listed firms from 1993-2010 reveals sample bias towards more profitable firms (see Appendix 3.8). Average market capitalisation coverage varies between 30 percent (for *Age*) and 72 percent (for *RFB* proxies) of total CMM market capitalisation ($323,647bn; Appendix 3.9). Given that, for example, the *Age* sample only covers 7 percent of all CMM listed firms, but accounts for 30 percent of CMM market cap, reveals that the sub-samples tend to be biased towards larger firms.

[42] Rubin (1976) provides three categories of missing data depending on the mechanism causing it; (1) missing-completely at random (MCAR), (2) missing at random (MAR) and (3) missing not at random (MNAR). When observations of a variable are *missing completely at random*, the missing observations are a random subset of all observations (i.e. there are no systematic differences between the missing and observed values) and the missing and observed values will have similar distributions. *Missing at random* means there might be systematic differences between the missing and observed values, but these can be entirely explained by other observed variables. For example, if analyst forecast precision data are missing at random, conditional on how long a firm has been listed on an exchange, then the distributions of missing and observed analyst forecast precision will be similar among firms with the same years of listing. Differently stated, if two firms share some observed values (e.g. same years of listing) they should also share similar statistical behaviour on other observations (e.g. analyst forecast precision), irrespective of whether those observations are observed or missing. In contrast, data is assumed to be *not missing at random* if neither MCAR nor MAR holds; that is, even after controlling for all the available observed data, the missing values remain related to the unseen observations themselves (Little and Rubin, 2002).

Table 3.2: Sample Characteristics by Constructs & Indicators

	Quantity			Precision			Asymmetry		
Fiscal Year	**Listing**	**Age**	**RMC**	**AQ**	**VR**	**AFP_Total**	**PIN**	**SPREAD**	**INV_Conc**
1993	1,038	574	760	1,019	2,215	523	1,866	2,847	1,845
1994	1,079	618	787	1,609	2,237	587	1,955	2,876	1,979
1995	1,099	669	789	1,667	2,198	577	2,089	2,841	2,002
1996	1,183	732	852	1,655	2,200	642	2,264	2,871	2,065
1997	1,310	847	933	1,665	2,171	670	2,678	2,941	2,110
1998	1,452	960	1,049	1,648	2,341	790	2,930	3,060	2,211
1999	1,533	1,041	1,078	1,681	2,438	848	2,992	3,073	2,324
2000	1,566	1,112	1,123	1,734	2,456	849	2,916	2,740	2,326
2001	1,685	1,207	1,220	1,852	2,540	942	3,033	3,295	2,505
2002	1,796	1,287	1,304	2,100	2,751	1,018	3,203	3,456	2,679
2003	1,878	1,368	1,369	2,179	2,841	1,151	3,357	3,511	2,763
2004	1,953	1,458	1,445	2,147	2,912	1,327	3,434	3,574	2,723
2005	1,987	1,531	1,523	2,233	2,975	1,466	3,444	3,568	2,813
2006	1,964	1,592	1,556	2,127	2,937	1,467	3,313	3,459	2,768
2007	1,969	1,667	1,588	2,090	2,817	1,501	2,139	3,389	2,750
2008	2,022	1,753	1,611	2,060	2,793	1,578	1,971	3,417	2,801
2009	2,128	1,869	1,680	2,056	2,789	1,599	3,297	3,502	2,897
2010	2,220	1,983	1,752	2,046	2,822	1,655	498	3,536	2,942
Total Obs.	**29,862**	**22,268**	**22,419**	**33,568**	**46,433**	**19,190**	**47,379**	**57,956**	**44,503**
Average Obs.	**1,659**	**1,237**	**1,246**	**1,865**	**2,580**	**1,066**	**2,632**	**3,220**	**2,472**
# of Firms	**3,295**	**1,987**	**1,887**	**4,042**	**5,486**	**3,618**	**6,288**	**6,907**	**6,167**
Avg. ROA	**3.75**	**4.76**	**4.73**	**4.30**	**3.28**	**3.96**	**3.46**	**3.06**	**3.39**
Avg. Mkt. Cap %	**31.58**	**29.76**	**36.41**	**46.81**	**61.45**	**35.11**	**61.04**	**68.69**	**55.12**

Table continued on next page.

Table 3.2: Sample Characteristics by Constructs & Indicators (cont.)

	Risk Factor-Based (RFB) CoE			Risk Factor-Based + VIX (FVIX) CoE			Valuation Model-Based (VMB) CoE		
Fiscal Year	**rCAPM**	**rFF3**	**rFF4**	**rFVIX**	**rFVIX3**	**rFVIX4**	**rPE**	**rPEG**	**rAEGM**
1993	2,920	2,920	2,920	2,920	2,920	2,911	1,543	1,129	1,125
1994	2,945	2,945	2,945	2,945	2,945	2,936	1,605	1,089	1,084
1995	2,923	2,923	2,923	2,923	2,923	2,915	1,595	1,135	1,116
1996	2,998	2,998	2,998	2,998	2,998	2,990	1,678	1,252	1,238
1997	3,094	3,094	3,094	3,094	3,094	3,086	1,724	1,347	1,331
1998	3,262	3,262	3,262	3,261	3,261	3,250	1,785	1,419	1,413
1999	3,333	3,333	3,333	3,332	3,332	3,321	1,754	1,317	1,309
2000	3,283	3,283	3,283	3,282	3,282	3,271	1,697	1,253	1,242
2001	3,450	3,450	3,450	3,449	3,449	3,438	1,708	1,555	1,545
2002	3,619	3,619	3,619	3,618	3,618	3,607	1,875	1,659	1,654
2003	3,692	3,692	3,692	3,691	3,691	3,680	2,069	1,901	1,897
2004	3,750	3,750	3,750	3,749	3,749	3,739	2,170	2,029	2,023
2005	3,739	3,739	3,739	3,738	3,738	3,728	2,217	2,053	2,043
2006	3,601	3,601	3,601	3,599	3,599	3,589	2,160	2,055	2,030
2007	3,519	3,519	3,519	3,516	3,516	3,507	2,074	2,017	1,992
2008	3,552	3,552	3,552	3,549	3,549	3,540	1,943	2,037	2,028
2009	3,639	3,639	3,639	3,636	3,636	3,627	2,130	2,253	2,235
2010	3,676	3,676	3,676	3,673	3,673	3,663	2,193	2,238	2,228
Total Obs.	**60,995**	**60,995**	**60,995**	**60,973**	**60,973**	**60,798**	**33,920**	**29,738**	**29,533**
Average Obs.	**3,389**	**3,389**	**3,389**	**3,387**	**3,387**	**3,378**	**1,884**	**1,652**	**1,641**
# of Firms	**7,091**	**7,091**	**7,091**	**7,088**	**7,088**	**7,074**	**4,998**	**4,933**	**4,919**
Avg. ROA	**3.06**	**3.06**	**3.06**	**3.06**	**3.06**	**3.05**	**4.08**	**3.82**	**3.83**
Avg. Mkt. Cap %	**71.86**	**71.86**	**71.86**	**71.86**	**71.86**	**71.21**	**70.43**	**55.92**	**51.97**

Average Return-On-Assets (Avg. ROA) equals the average median ROA (Income Before Extraordinary Items / Total Assets) from 1993 to 2010. Average Market Capitalisation (Avg. Mkt. Cap %) equals the average market capitalisation (common shares outstanding x fiscal year end closing stock price) in percent of total CMM market cap ($323,647bn) from 1993 to 2010. Detailed ROA and Market Capitalisation figures by constructs and indicators are provided in Appendix 3.8 and Appendix 3.9, respectively. Variable names as defined in Table 3.1.

3.3.3 SEM Estimation and Model Fit Statistics

Before estimating the SEM models, I standardise all indicator variables to mean zero and variance one to ensure equal contribution to their respective scale. All models are run in MPLUS and converged to an admissible solution. The method of estimation is full information maximum likelihood estimation with standard errors adjusted for firm clusters. I only report standardised model estimates since unstandardized parameter coefficients are of no economic meaning. When comparing models between groups, I assume measurement equality and apply chi-square difference tests to examine whether differences in structure coefficients across the groups are statistically significant (Kline, 2011, p. 215).[43]

Model fit is assessed on how well a particular SEM specification resembles *observed* variance-covariance in the sample data with *predicted* variance-covariance by the model; that is, the better actual correlations for the empirical measures of *Quantity*, *Precision*, *Asymmetry* and *CoE* reconcile with estimated ones by the conceptual SEM, the greater overall model fit. Given the inapplicability of the chi-square statistic in large sample studies like this one, I follow Hu and Bentler (1999) and appraise model fit according to the following index levels: (1) TLI ≥ 0.95 and SRMR ≤ 0.09; (2) CFI ≥ 0.95 and SRMR ≤ 0.09; and (3) RMSEA ≤ 0.06 and SRMR ≤ 0.09.[44] In addition, I examine correlation residuals (calculated as the difference between model-implied and observed-sample correlations) to detect sample correlations that are not well explained by the overall model. As a general rule, absolute differences greater than 0.10 are regarded as problematic (Kline, 2011, p. 171).

---

[43] To test for statistical significant differences in structure coefficients, I compare the model chi-square statistic of the nested model (at which the structural path of interest is constraint to be equal between groups), with the comparison model where all paths are free to vary. If the Chi-square difference is statistically significant, the equal-fit hypothesis is rejected (i.e., the difference between group coefficients is statistically significant).

[44] The model chi-square statistic is the most commonly used model test statistic in SEM research. It is structured as a "badness-of-fit" measure in that higher values indicate worse model fit: significant results (say, $p < 0.05$) denote overall model misspecification. However, unless perfect model fit is attained– which is unlikely in any real world application – the model chi-square statistics increases with sample size. Thus, in very large samples (e.g. N = 5,000, Kline, 2011, p. 201) even minor model-data discrepancies can lead to test statistics rejecting an otherwise valid model (Fan et al., 1999). With number of observations well exceeding this threshold (N = 60,995 for most of my models), the model chi square statistic is an invalid measure of model fit in this study.

3.4 Results

3.4.1 Descriptive Statistics

Table 3.3 and Table 3.4 show descriptive statistics and Pearson correlations for the indicators. In the text, I also report for each construct its Cronbach's alpha that, on a scale from zero to one, provides guidance on how well the empirical proxies used in this study measure *Quantity*, *Asymmetry*, *Precision* and *CoE*. Coefficients below 0.50 indicate that most observed score variance is due to random error. However, when sample size is sufficiently large—as it is in this paper—somewhat lower Cronbach's alphas are acceptable (Kline (2011), Little et al. (1999)).

3.4.1.1 Information Quantity

The average (median) firm in my sample is incorporated for 27.8 (18.0) years and listed on one of the three major US exchanges for 14.7 (11.0) years, implying that the average firm is in existence for about 13 (seven) years before going public. These figures are similar to median IPO age figures reported on Jay Ritter's website.[45] A negative median value for the *RMC* index suggests that firms in the sample tend to enjoy less media coverage than the average company in the same industry and of similar size. Because this is a novel approach to measure media coverage, comparability with other studies is not possible. All three indicators are significantly positively correlated with an adequate Cronbach's alpha of 0.66.[46]

3.4.1.2 Information Precision

Sample characteristics for the earnings quality indicators are consistent with previous studies, but less so with respect to total analyst forecast precision (*AFP_Total*). An average (median) *AQ* metric of -0.061 (-0.040) for my sample compares well with FLOS (2005), who report a mean of 0.044 (0.031). My average *VR* metric of -0.410 (-0.348) is similar to figures reported in FLOS (2004)—mean: 0.423 (0.416).[47] However, mean

[45] https://site.warrington.ufl.edu/ritter/ipo-data/
[46] In line with my main SEM results, Cronbach's alphas are based on standardised indicator variables.
[47] As noted above, each of the earnings quality metrics is multiplied by negative one which explains the differing signs between my metrics and those of previous studies.

(1,059), median (947) and interquartile range (1,310) figures for *AFP_Total* are inconsistent with statistics reported in Barron et al. (2012)—mean: 3,049, median: 288, IQR: 1,312—and Botosan and Plumlee (2013)—mean: 2,113, median: 947, IQR: 2,847—which signifies an increased sensitivity of this measure towards extreme outliers and measurement error (Barron et al., 2012, p. 21).

Although all three indicators are significantly positively correlated, the economic significance of the association between total analyst forecast precision and the two earnings quality metrics is negligible (ρ: 0.019 and 0.020). The reduced relevance of *AFP_Total* is also reflected by a low Cronbach's alpha of 0.43, which only increases to an acceptable level of 0.54 once total analyst forecast precision is excluded from the measurement model. Taken together, this points in the direction of *AFP_Total* not measuring information precision to the extent it was hoped.

3.4.1.3 Information Asymmetry

Descriptive statistics for both market microstructure and market competition indicators are matching results in previous studies. A mean *PIN* score of 0.205 for the average firm in my sample falls within the range of reported average scores of 0.150 to 0.300 in extant work (Bhattacharya et al. (2012), Brown and Hillegeist (2007), Duarte et al. (2008); Easley et al. (2002)) and sample *SPREADS* (mean: 0.019; median: 0.010) are compatible with several differently estimated spread figures in Corwin and Schultz (2012, Table 3). Average *INV_Conc* (mean 0.195; median: 0.103) is like absolute figures reported in Akins et al. (mean: -0.17; median: -0.09) and all three indicators are significantly positively correlated with an adequate Cronbach's alpha of 0.65.[48]

3.4.1.4 Cost of Equity and Future Realised Returns

Summary statistics for the VMB and RFB proxies are widely consistent with former evidence. Reported mean (0.124) and median (0.103) values for *rPEG* are comparable to results in Barron et al. (2012)and Easton and Monahan (2005). Similarly, mean (0.103) and median (0.093) expected returns for *rFF4* are like figures in Barth et al. (2013) and

---

48 Opposing sings explained by the fact that Akins et al. multiply the index by negative one, from which I refrain to maintain consistency among the indicators.

Kothari et al. (2009).[49] The VIX augmented RFB proxies *rFVIX*, *rFVIX3* and *rFVIX4* are equal to their VIX-free counterparts in terms of average predicted CoE levels, but with somewhat higher standard deviations.

Consistent with findings in the comparative literature (Lee et al. (2015), Lee et al. (2010)), VMB and RFB measures are slightly negatively correlated. This suggests that they "do not capture the same underlying construct" (Botosan et al., 2011, p. 1102).[50] However, if there is no clear theory (or, as in this case, an ongoing debate) about which indicators are more valid estimates of a particular construct, the inclusion of all available, rather than only highly homogenous indicators, is recommended, even though they might be plagued by low internal consistency (Little et al., 1999, p. 207). This further supports the use of both VMB (*rPEG*) and RFB (*rFF4*) proxies as indicators for the CoE construct in the first part of this analysis.

Average one-year ahead realised returns (17.1 percent) are similar to findings in previous work, but median future realised returns (8.7) are somewhat lower (e.g., Botosan et al. (2011), Guay et al. (2011)). Average annualised two- and three-year ahead returns (10.8 and 8.9) converge towards median levels of 8.6 and 7.8 percent over time. Among the VMB proxies, *rPEG* exhibits the strongest positive association with FRR (ret12: 0.148; ret24: 0.121; ret36: 0.097); in contrast, *rFF4* is best-behaved among the RFB proxies in that it shows the least negative association with FRR (ret12: -0.022; ret24: -0.074; ret36: -0.089).[51] These findings conform to the general tenor in the performance literature of VMB proxies being more reliable measures of CoE than RFB proxies, given their stronger positive association with future realised returns (Lee et al., 2015).

---

[49] In contrast to this paper, both studies manipulate CoE to lie between 0.00 and 0.50 which explains slightly higher mean and median values then in this paper (e.g. mean in Barth et al.: 0.160; Kothari et al.: 0.146). However, if setting all observation to the same range, my statistics are similar (mean: 0.134).

[50] Cronbach's alpha not applicable due to negative correlations between the proxies.

[51] The inclusion of (*ex ante*) investor sentiment about future market volatility in the CoE estimates only slightly improves the negative correlations with future returns. For instance, *rFVIX4*, the volatility risk factor augmented *rFF4* proxy, shows significant negative correlations of -0.021, -0.066 and -0.083 with annualised 12-, 24- and 36-month buy-and-hold returns.

Table 3.3: Summary Statistics by Constructs & Indicators

Construct Indicator	**Mean**	**Std. Dev.**	**10%**	**25%**	**Median**	**75%**	**90%**
Quantity							
Listing	14.668	12.614	5.000	7.000	11.000	17.000	27.000
Age	27.862	26.232	7.000	11.000	18.000	35.000	70.000
RMC	-10.134	1663.926	-816.942	-535.663	-245.481	111.810	591.685
Precision							
AQ	-0.061	0.067	-0.128	-0.073	-0.040	-0.023	-0.014
VR	-0.410	0.269	-0.749	-0.523	-0.348	-0.225	-0.145
AFP_Total	1059.000	1440.074	29.941	105.310	412.833	1414.583	3070.200
Asymmetry							
PIN	0.205	0.107	0.088	0.123	0.184	0.268	0.352
SPREAD	0.019	0.027	0.001	0.003	0.010	0.024	0.046
INV_Conc	0.195	0.224	0.029	0.048	0.103	0.248	0.508
CoE-RFB							
rCAPM	0.084	0.106	-0.018	0.006	0.063	0.138	0.222
rFF3	0.115	0.127	-0.009	0.038	0.104	0.181	0.262
rFF4	0.103	0.208	-0.026	0.026	0.093	0.174	0.261
CoE-FVIX							
rFVIX	0.079	0.127	-0.052	0.005	0.063	0.140	0.230
rFVIX3	0.115	0.148	-0.034	0.040	0.110	0.187	0.273
rFVIX4	0.102	0.238	-0.057	0.022	0.100	0.184	0.276
CoE-VMB							
rPE	0.069	0.060	0.030	0.047	0.064	0.084	0.109
rPEG	0.124	0.096	0.068	0.084	0.103	0.137	0.196
rAEGM	0.139	0.310	0.079	0.093	0.113	0.146	0.206
CoE-FRR							
ret12	0.171	1.169	-0.361	-0.125	0.087	0.324	0.659
ret24	0.108	0.356	-0.253	-0.072	0.086	0.242	0.455
ret36	0.089	0.251	-0.193	-0.045	0.078	0.204	0.365

The table reports summary statistics for the indicator variables used in the SEM analysis. Variable names as defined in Table 3.1.

Table 3.4: Pearson Correlations for Indicators

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
(1)	Listing	**1.000**																				
(2)	Age	.611	**1.000**																			
(3)	RMC	.228	.119	**1.000**																		
(4)	AQ	.188	.196	-.031	**1.000**																	
(5)	VR	.135	.181	-.022	.370	**1.000**																
(6)	AFP_Total	-.013†	.048	-.022	.019	.020	**1.000**															
(7)	PIN	-.207	-.150	-.068	-.074	-.032	.086	**1.000**														
(8)	SPREAD	-.156	-.058	.036	-.071	-.096	.058	.420	**1.000**													
(9)	INV_Conc	-.168	-.143	.011†	-.123	-.094	.029	.522	.491	**1.000**												
(10)	rCAPM	-.058	.001†	-.008†	-.040	-.021	.102	-.070	.133	.008†	**1.000**											
(11)	rFF3	-.049	.000†	-.060	.030	.018	.056	-.024	.076	-.007†	.597	**1.000**										
(12)	rFF4	-.031	-.001†	-.049	.032	.049	.058	-.038	.015	-.015	.258	.745	**1.000**									
(13)	rFVIX	-.034	.029	.003†	.014	.052	.099	-.065	.137	.013	.928	.563	.243	**1.000**								
(14)	rFVIX3	-.030	.008†	-.055	.023	.032	.051	-.046	.018	-.035	.481	.865	.653	.521	**1.000**							
(15)	rFVIX4	-.018	.009†	-.049	.028	.077	.056	-.059	-.022	-.029	.215	.683	.944	.246	.744	**1.000**						
(16)	rPE	.000†	.022	-.010†	.094	.064	-.108	.127	.150	.142	-.007†	.011	-.005†	.011	-.004†	-.009	**1.000**					
(17)	rPEG	-.045	-.057	.000†	-.186	-.205	-.176	.135	.179	.129	-.042	-.053	-.057	-.061	-.052	-.058	.140	**1.000**				
(18)	rAEGM	-.011†	-.012†	.000†	-.064	-.070	-.106	.045	.081	.048	-.001†	-.007†	-.015	-.005†	-.009†	-.016	.061	.309	**1.000**			
(19)	ret12	-.021	-.021	-.005†	-.012	-.035	.011†	.049	.104	.055	-.054	-.025	-.022	-.047	-.017	-.021	.076	.148	.033	**1.000**		
(20)	ret24	-.016	-.010†	-.009†	.017	-.027	.038	.091	.123	.077	-.150	-.099	-.074	-.136	-.082	-.066	.089	.121	.023	.447	**1.000**	
(21)	ret36	-.003†	.000†	-.013	.047	-.002†	.044	.101	.123	.077	-.138	-.102	-.089	-.124	-.085	-.083	.086	.097	.019	.320	.777	**1.000**

*Pearson correlation coefficients based on pairwise deletion. † p > 0.10, * p > 0.05. All other coefficients are significant at the 5% level or better. Variable names as defined in Table 3.1.*

3.4.2 Results of SEM Analyses

Table 3.5 reports main results along with fit and descriptive statistics for all SEM models tested in this paper. Based on these results, I provide several figures to ease interpretation of my findings.

3.4.2.1 Measurement Model

Before the direct and indirect links between the information attributes and CoE can be analysed, an acceptable measurement model needs to be attained. Model fit must be adequate, and each attribute should load statistically significantly with predicted signs on its pre-specified indicators. As a base model, I first estimate the conjectured SEM model and report factor loadings and structure coefficients in Figure 3.4. Fit statistics for my base model show good levels of fit for the structural model (SRMR: 0.048) and acceptable levels for the measurement model (CFI: 0.912; TLI: 0.873; RMSEA: 0.013, Table 3.5, Panel C).

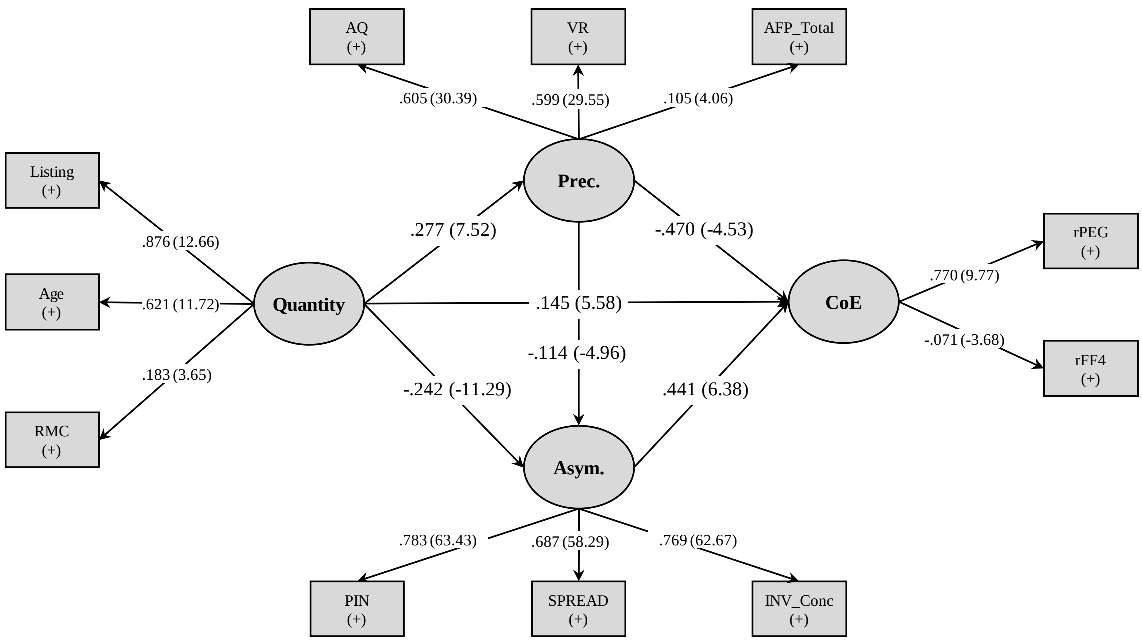


Figure 3.4: Base Model

All coefficients standardized and significant at the 1% level or better. T-statistics based on standard errors clustered by firm are reported in parentheses. Error terms are suppressed. Variable names as defined in Table 3.1.

As conjectured, all proxies are positively correlated with their respective information attribute. *Quantity* loads significantly positively on all its indicators with most variance explained in *Listing* [$R^2 = (0.876)^2 = 0.767$]. Factor loadings on *PIN*, *SPREAD* and *INV_Conc* are all positive and significant, with *Asymmetry* explaining on average 56 percent of indicators' variance. *Precision* is substantially correlated with the earnings quality indicators *AQ* and *VR*, but only weakly associated with *AFP_Total* ($R^2 = 0.011$). Interestingly, the CoE construct explains about 60 percent of variance in *rPEG*, but practically none in *rFF4*. This indicates that VMB proxies are more likely to capture changes in firms' information environments than RFB proxies. I elaborate further on this point in Section 3.4.3 Results of CoE Analyses. As a result, the measurement of *CoE* is less well behaved; however, I follow the recommendation of Little et al. (1999) and keep both *rPEG* and *rFF4* as CoE proxies.

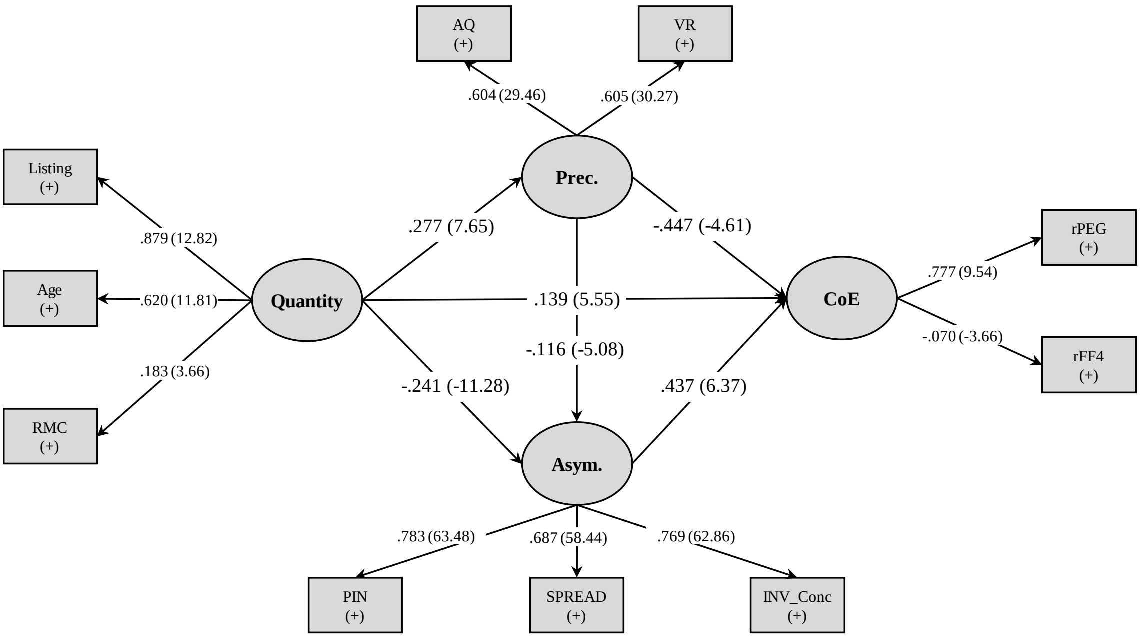


Figure 3.5: Best-Fit Model

All coefficients standardized and significant at the 1% level or better. T-statistics based on standard errors clustered by firm are reported in parentheses. Variable names as defined in Table 3.1.

Closer inspection of the correlation residuals of the base model reveals that four residuals exceed the critical value of 0.10; *AFP_Total* is responsible for three of these violations (see Table 3.6). In particular, the model over-predicts its association with *PIN* (correlation residual: 0.148) and *SPREAD* (0.140), but under-predicts its association with *rPEG* (-0.138). Excluding *AFP_Total* from the base model leaves factor loadings and structure coefficient almost unchanged (see Figure 3.5), but increases overall model fit considerably. This is supported by a significant chi-square difference between the two models (chi-square difference: 265.8; p-value: 0.000) and improved fit indexes for the best-fit model (Table 3.5, Panel C). Thus, all subsequent analyses are based on this best-fit SEM specification.

To summarise, information asymmetry between investors is reliably measured by all three empirical proxies (*PIN*, *SPEARD*, *INV_Conc*); similarly, the amount of available information to investors is well measured by both the period of time a firm is listed (*Listing*) and in existence (*Age*), with media coverage contributing significantly, but economically weakly, to its measurement (*RMC*). Notably, firms' information precision tends to be best proxied by earnings quality metrics (*AQ*, *VR*); that is, exclusion of *AFP_Total* increases model fit substantially. One explanation might be that *AFP_Total* is a catch-all measure for the overall quality of firms' information environments, rather than a proxy for *Precision*: analysts take a holistic approach in appraising companies, and thus, the accuracy of their forecasts not only depends on the precision of provided information, but also (1) on the extent to which they have access to private information and (2) on the amount of firm-specific information available to them. Therefore, if one seeks an aggregate measure for the overall quality of a firm's information environment, *AFP_Total* seems to be a valid proxy as substantiated by its significant negative association with all three VMB proxies (*rPE*: -0.108; *rPEG*: -0.176; *rAEGM*: -0.106; see Table 3.4). If, on the other hand, a more detailed examination of the respective information attributes is required, their measurement tends to be best conducted by the proxies suggested in the best-fit model.

Table 3.5: SEM Model Variations

Panel A: Structural Coefficients (standardized)

						Grouping: Exchange			Grouping: Competition		Grouping: Time	
Direct Effects			Base	Best-Fit	BEOS	AMEX	NASDAQ	NYSE	LOW	HIGH	1993/99	2000/10
Quant.	→	Prec.	0.277	0.277	0.265	0.276†	0.260	0.217	0.175	0.318	0.241	0.259
Quant.	→	Asym.	-0.242	-0.241	-0.251	-0.250†	-0.272	-0.303	-0.122	-0.155	-0.210	-0.108
Prec.	→	Asym.	-0.114	-0.116	-0.083	-0.009†	0.122	-0.275	0.060	-0.057	-0.238	-0.171
Quant.	→	CoE	0.145	0.139	0.134	0.206†	0.185	0.096	0.104	0.086	0.048*	0.128
Asym.	→	CoE	0.441	0.437	0.572	0.490*	0.450	0.206	0.209	0.058	0.372	0.667
Prec.	→	CoE	-0.470	-0.447	-0.697	-0.629	-0.454	-0.334	-0.519	-0.328	-0.428	-0.481
Indirect & Total Effects			Base	Best-Fit	BEOS	AMEX	NASDAQ	NYSE	LOW	HIGH	1993/99	2000/10
Quant. (direct)	→	CoE	0.145	0.139	0.134	0.206†	0.185	0.096	0.104	0.086	0.048*	0.128
Quant. (ind.)	→	CoE	-0.251	-0.243	-0.341	-0.279†	-0.226	-0.147	-0.114	-0.114	-0.202	-0.226
Quant (total)	→	CoE	-0.105	-0.104	-0.207	-0.091†	-0.042	-0.051	-0.011†	-0.028†	-0.154	-0.098
Prec. (direct)	→	CoE	-0.470	-0.447	-0.697	-0.629	-0.454	-0.334	-0.519	-0.328	-0.428	-0.481
Prec. (ind.)	→	CoE	-0.05	-0.051	-0.047	-0.004†	0.055	-0.056	0.012†	-0.003	-0.089	-0.114
Prec. (total)	→	CoE	-0.52	-0.498	-0.745	-0.633	-0.399	-0.391	-0.507	-0.331	-0.512	-0.595

Table continued on next page.

Table 3.5: SEM Model Variations (cont.)

Panel B: Explained Variance (R-Square)

				Grouping: Exchange			Grouping: Competition		Grouping: Time	
Dependent Variable	Base	Best-Fit	BEOS	AMEX	NASDAQ	NYSE	LOW	HIGH	1993/99	2000/10
CoE	0.414	0.413	0.860	0.603	0.339	0.182	0.292	0.103	0.396	0.761
Prec.	0.076	0.077	0.070	0.076	0.068	0.047	0.031	0.101	0.058	0.067
Asym.	0.087	0.087	0.081	0.064	0.072	0.204	0.016	0.033	0.124	0.050

Panel C: Fit Statistics

				Grouping: Exchange			Grouping: Competition		Grouping: Time	
Statistic	Base	Best-Fit	BEOS	AMEX	NASDAQ	NYSE	LOW	HIGH	1993/99	2000/10
DoF.	38	29	29	111	111	111	70	70	70	70
χ^2	413.2	147.4	607.2	906.3	906.3	906.3	2332.8	2332.8	730.2	730.2
χ^2 p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CFI	0.912	0.959	0.848	0.833	0.833	0.833	0.607	0.607	0.840	0.840
TLI	0.873	0.936	0.764	0.797	0.797	0.797	0.495	0.495	0.794	0.794
RMSEA	0.013	0.008	0.022	0.019	0.019	0.019	0.038	0.038	0.018	0.018
SRMR	0.048	0.038	0.050	0.072	0.072	0.072	0.090	0.090	0.090	0.090
Obs.	60,995	60,995	43,008	4,018	27,612	29,365	22,251	22,252	21,475	39,520

Table continued on next page.

Table 3.5: SEM Model Variations (cont.)

Panel D: Descriptive Statistics										
				Grouping: Exchange			**Grouping: Competition**		**Grouping: Time**	
Mean	Base	Best-Fit	BEOS	AMEX	NASDAQ	NYSE	LOW	HIGH	1993/99	2000/10
Age (years)	27.86	27.86	27.94	28.36	23.81	30.55	26.57	33.36	29.08	27.47
Listing (years)	14.67	14.67	13.64	12.31	11.75	17.89	12.42	17.77	13.00	15.35
Market Cap (mUSD)	3,861.29	3,861.29	3,314.54	310.50	1,494.65	6,569.55	1,166.39	7,213.98	2,656.83	4,515.22
ROA (%)	1.85	1.85	1.40	-1.47	-2.50	7.19	3.04	3.89	2.75	1.35

*† p > 0.10, * p > 0.05. All other coefficients are significant at the 5% level or better. BEOS model based on sample 1993 – 2005. Competition groups formed on basis of INV_Conc and split at the 50th percentile. Measurement equality across groups is assumed.*

Table 3.6: Correlation Residuals for Hypothesised (Base) SEM Model

	Listing	Age	RMC	AQ	VR	AFP-Total	PIN	SPR-EAD	INV-Conc	rPEG	rFF4
Listing	**0.005**										
Age	-0.003	**-0.002**									
RMC	0.022	-0.037	**0.003**								
AQ	0.001	0.079	-0.073	**0.006**							
VR	-0.023	0.062	-0.070	0.006	**0.004**						
AFP_Total	-0.055	0.017	-0.023	-0.043	-0.041	**0.031**					
PIN	-0.033	0.020	-0.005	0.040	0.055	0.148	**0.013**				
SPREAD	0.001	0.066	0.086	-0.002	-0.021	0.140	0.014	**0.003**			
INV_Conc	0.014	-0.007	0.095	-0.101	-0.072	0.081	0.007	-0.001	**0.001**		
rPEG	-0.001	-0.020	0.049	0.021	0.001	-0.138	-0.056	0.078	-0.071	**0.010**	
rFF4	-0.066	-0.017	-0.070	0.009	0.034	0.064	-0.008	0.070	0.003	0.002	**0.000**

Correlation residuals are the difference between model-implied correlations and observed (sample) correlations.

3.4.2.2 Structural Model

Having established the measurement model (which as a corollary provides novel evidence on the relative importance of the various information proxies used in extant research), the focus of our analysis can now be shifted towards the first research question: To what extent does information quantity, precision and asymmetry affect a firm's CoE?

Consistent with hypotheses H4 and H5, the best-fit model (see Figure 3.5) shows a significant positive association between *Quantity* and *Precision* (0.277; p-value: 0.000) and a significant negative association between *Quantity* and *Asymmetry* (-0.242; p-value: 0.000). This endorses the notion that information quantity is the precursor or antecedent of these two attributes. Only as the amount of available information about a firm increases (e.g., increased disclosure) can investors begin to evaluate the accuracy of the information and informational disadvantages between the informed and uninformed can be resolved (i.e., asymmetry decreases). The direct effect of *Quantity* on *CoE* diverges from its assumed negative association (0.139; p-value: 0.000; rejecting H3); however, the total effect remains significantly negative (-0.104; p-value: 0.000), given that the indirect effects (i.e., *Quantity* on *CoE* through *Precision* and *Asymmetry*, respectively) are almost two and a half times as important (-0.243; p-value: 0.000) than the direct one.

With respect to the direct associations between *Precision*, *Asymmetry* and *CoE*, it is shown that as firms increase the accuracy of provided information (e.g., by increased earnings quality), they can expect to enjoy favourable CoE effects (-0.447; p-value: 0.000; confirming H1). Similarly, firms that enact effective measures to reduce information asymmetry between the privately informed and publicly uninformed (e.g., decrease investor concentration) tend to benefit from lower return expectations (0.437; p-value: 0.000; confirming H2). The direct *Asymmetry* and *Precision* effects in this study are about twice as strong as those reported in BEOS (2012) and are even stronger when the model is re-estimated over their sampling period from 1993 to 2005 (see BEOS model in Table 3.5). The fact that the average firm in BEOS (2012) is larger and more established than in my study and, furthermore, that information risk tends to be conditional on firm size and maturity (e.g., Clarkson and Thompson (1990); Ecker (2014)), may explain these magnified effects.[52]

Consistent with BEOS (2012), the direct effect of *Precision* on *CoE* accounts for almost 90 percent of the total effect between the two constructs (-0.498; p-value: 0.000), with the indirect effect explaining 10 percent of the total association (*Precision* via *Asymmetry* on *CoE:* -0.051 = -0.116 x 0.437). The rather low importance of the indirect effect can be explained by the weak negative association between *Precision* and *Asymmetry* (-0.116; p-value: 0.000; confirming H3). Given this finding, BEOS concludes that “if there is a trade-off between improving the quality or precision of information and increasing equality of access to information [decreasing *Asymmetry*], our results suggest that the former effect dominates the latter effect” (p. 477).

However, their conclusion hinges upon a weak economic association between *Asymmetry* and *Precision*. For instance, if one assumes that lower *Asymmetry* causes higher *Precision*, instead of higher *Precision* causing lower *Asymmetry* (i.e., interchanging the causality assumption in our SEM model), the conclusion changes accordingly. In such a model (unreported), all structure coefficients are unchanged, but because the indirect effect is mediated by information precision—*Asymmetry* via *Precision* on *CoE* (indirect effect: 0.051 = -0.115 x -0.447)—the direct effect of *Asymmetry* (and not *Precision*) accounts now for 90 percent of the total CoE effect. Therefore, the debate is arguably less a question of which information attribute dominates the other in terms of its impact on *total* CoE effects than which attribute do firms have greater discretion over? Given that both direct effects show similar strength, I posit that firms can enjoy equally strong (*direct*) positive CoE effects from either providing more accurate information to investors or distributing private information more equally between them. Thus, if there is a trade-

[52] The BEOS sample only contains Value Line firms that have existed for at least seven years, resulting in a total sample of 12,648 firm-years, for which they report an average sample ROA of 5.06 percent and market cap coverage of 42.96 percent. Over the same period, my IBES sub-sample contains 19,138 firm-years, with ROA of 3.84 percent and market cap coverage of 50.81 percent, illustrating that my sample contains a greater number of smaller, younger and less profitable firms than in BEOS. Market cap coverage equals percentage of total market capitalization accounted for by sample firms; BEOS uses CRSP figures, this study uses CMM figures for total market cap. Difference between Value Line and IBES analyst forecasts discussed in Philbrick and Ricks (1991).

off, firms should dedicate scarce resources to those activities for which the potential of improvement is expected to be most positive.

To summarise, these initial results corroborate the proposition that firm-specific information has a significant impact on firms' CoE on its merits, but also show that firms' information environments are clearly governed by the influence of *Asymmetry* and *Precision*, which raises the question of what relevance *Quantity* is to firms and investors alike. The next section elaborates on this.

3.4.2.3 Relative Importance of Idiosyncratic Information

A central objective of this study is to offer new insights on the extent to which differences in firms' information environments affect investors' return expectations. To assess the relative importance of idiosyncratic information as a determinant of CoE, I test for systematic differences between (1) firms with different firm-characteristics; (2) firms with different levels of market competition; and (3) sub-sample periods 1993/99 and 2000/10. Table 3.7 reports results of these tests.

Firm Characteristics. As a natural experiment, I categorise the sample firms into groups of stock exchange membership to test for structural differences between firms of different size, age and profitability. As shown in Figure 3.6, the average NYSE firm is much larger (market cap: 6,569 mUSD), older (31 years), longer listed (18 years) and more profitable (ROA: 7.2 percent) than firms trading on AMEX and NASDAQ.[53] The rather low structure coefficients between *Quantity* and *Precision* (0.277) and *Quantity* and *Asymmetry* (-0.241) expose some sort of a ceiling effect for information quantity. That is, while *Quantity* tends to be of increased importance to investors when they are assessing the prospects of smaller, younger and less profitable firms, its relevance appears to plateau as firms mature. Evidence for this supposition is provided by decreasing direct, indirect and total effects of *Quantity* on *CoE* as one moves from AMEX (direct: 0.206; indirect: -0.279; total: -0.091) to NASDAQ (0.185; -0.226; -0.042) to NYSE stocks (0.096; -0.147; -0.051). This observation is consistent with findings in the estimation risk literature which

[53] Unreported one-way ANOVA tests show that all mean levels for *Age*, *Listing*, *Market Cap* and *ROA* are significantly different between exchanges (p-value: 0.000), except for *Listing* and *ROA* differences between AMEX and NASDAQ firms (p-value: 0.268 and 1.000).

shows that for stocks for which there is limited amount of information (i.e., smaller and younger firms), investors have difficulties accurately estimating their return parameters, which makes these firms more risky investments and hence induces higher CoE (e.g., Clarkson and Thompson (1990), Ecker (2014)).

What is more, NYSE firms are significantly less exposed to information risk than AMEX and NASDAQ stocks. First, this is delineated by significantly different direct effects between the exchanges: for instance, the direct impact of *Asymmetry* on *CoE* for NASDAQ (0.450, p-value: 0.000) and AMEX (0.490, p-value: 0.026) stocks is about twice as strong than for NYSE stocks (0.206, p-value: 0.001). Correspondingly, AMEX firms tend to benefit most (in the form of reduced CoE) from providing more precise information to investors (-0.629; p-value: 0.005), whereas these effects are lowest for NYSE (-0.334; p-value: 0.000) and moderate for NASDAQ stocks (-0.454; p-value: 0.000). Second, the model explains just 18.2 percent of CoE variance for NYSE firms, but a substantial 60.3 and 33.9 percent for AMEX and NASDAQ stocks, respectively.

Descriptive Statistics			
	AMEX	NASDAQ	NYSE
Age (yrs)	28.36	23.81	30.55
Listing (yrs)	12.31	11.75	17.89
Market Cap. (m$)	310.50	1,494.65	6,569.55
ROA (%)	-1.47	-2.50	7.19
Model Statistics			
CoE R^2	0.603	0.339	0.182
CFI	0.833	0.833	0.833
SRMR	0.072	0.072	0.072
Obs.	4,018	27,612	29,365

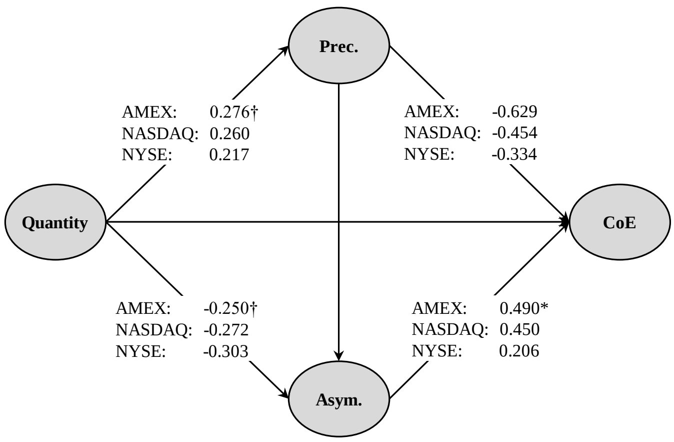


Figure 3.6: Best-Fit Model by Stock Exchanges

*The figure shows standardized structural coefficients of the best-fit model, categorised by firms' stock exchange membership (AMEX, NASDAQ, NYSE), along with descriptive and model statistics. Information extracted from Table 3.5. † p > 0.10, * p > 0.05. All other coefficients are significant at the 5% level or better.*

Table 3.7: Chi-square Difference Tests Between Structure Coefficients Across Groups

Panel A: Differences in Structural Coefficients (standardized)

			Grouping: Exchange						Grouping: Competition		Grouping: Time	
			AMEX – NYSE		**NASDAQ – NYSE**		**AMEX – NASDAQ**		**LOW – HIGH**		**1993/99 – 2000/10**	
Direct Effects			Difference	p-value	Difference	p-value	Difference	p-value	Difference	p-value	Difference	p-value
Quant.	→	Prec.	0.059*	0.100	0.043	0.000	0.016†	0.163	-0.143†	0.999	-0.018	0.002
Quant.	→	Asym.	0.053	0.026	0.031	0.000	0.022*	0.095	0.033	0.000	-0.102	0.000
Prec.	→	Asym.	0.266	0.000	0.397	0.000	-0.131*	0.095	0.117	0.007	-0.067	0.000
Quant.	→	CoE	0.110	0.008	0.089	0.000	0.021	0.010	0.018	0.000	-0.078	0.001
Asym.	→	CoE	0.284	0.009	0.244*	0.062	0.040	0.012	0.151	0.000	-0.295	0.000
Prec.	→	CoE	-0.295	0.004	-0.120†	0.153	0.295	0.004	-0.191	0.000	0.053	0.012

*† p > 0.10, * p > 0.05. All other coefficients are significant at the 5% level or better;* p*-value of 0.999 denotes a negative Chi-square difference between the nested and comparison model and as such a statistically insignificant difference. Competition groups formed on basis of INV_Conc and split at the 50th percentile.*

Market Competition. There is an ongoing debate if—as shown in LLV (2012)—*Asymmetry* effects are indeed conditional on market competition (Akins et al. (2012), Armstrong et al. (2011), Barron et al. (2012)). To evaluate this conjecture and to contribute new evidence to this discussion, I divide my sample into two groups of LOW and HIGH competition firms based on *INV_Conc* (where higher institutional investor concentration denotes lower competition) and test for significant differences between the structural coefficients across the two groups (all differences significant at the 1% level, except for *Quantity* on *CoE*, see Table 3.7). As Figure 3.7 shows, the direct effect of *Asymmetry* on *CoE* for the low competition group (0.209, p-value: 0.003) is about 3.5 times as strong as it is for the high competition group (0.058 p-value: 0.000) which suggests that as markets become more liquid (i.e., approaching perfect market competition), the impact of asymmetric information on CoE declines. Similarly, the effects of *Precision* on *CoE* are significantly higher for low competition firms than for the high competition segment. Considering the explanatory power of all three attributes, the model explains three times as much CoE variation for LOW group stocks than for HIGH group firms (R-Square: 29.2 and 10.3 percent, respectively).

These patterns are comparable to the dampened impact of idiosyncratic information on firms' expected returns as one moves from the smaller and younger AMEX/NASDAQ firms to the lager and older NYSE stocks. Given that the average firm in the LOW (HIGH) competition sample exhibits similar firm characteristics as the average AMEX/NASDAQ (NYSE) firm, this is unsurprising.[54] However, it illustrates that *Asymmetry* and *Precision* effects tend to move in concert with each other in that they simultaneously gain and lose in relevance as size and maturity characteristics of firms change. Differently stated, it is not only *Asymmetry*, but also *Precision* effects which decline as market competition increases and firms become older and more established: this contrasts with LLV's (2012) conjecture and can be seen as an empirical extension of their model. Thus, it might be fruitful to identify a common information risk factor that captures both firms' asymmetry and precision characteristic in order to bridge the gap between asymmetry/liquidity research (e.g., Amihud (2002), Amihud and Mendelson (1986)) and precision/accounting

[54] Unreported one-way ANOVA tests show that all mean levels for Age, Listing, Market Cap and ROA are significantly different between LOW and HIGH groups (p-value: 0.000).

quality research (e.g., FLOS (2004, 2005)). A first promising step in this direction is taken by Ecker et al. (2006).

Descriptive Statistics	LOW	HIGH
Age (yrs)	26.57	33.36
Listing (yrs)	12.42	17.77
Market Cap. (m$)	1,166.39	7,213.98
ROA (%)	3.04	3.89
Model Statistics		
CoE R^2	0.292	0.103
CFI	0.607	0.607
SRMR	0.090	0.090
Obs.	22,251	22,252

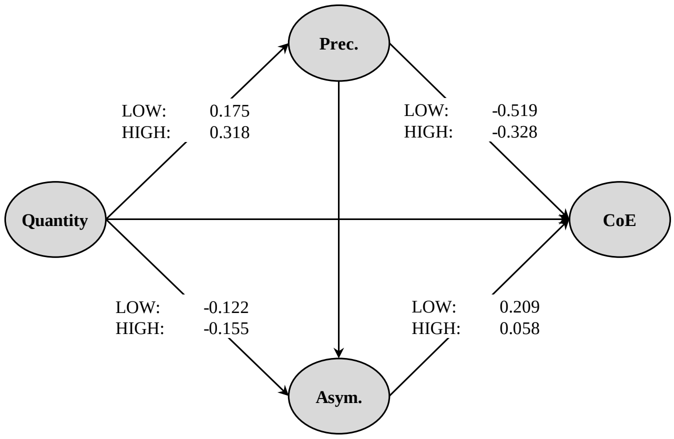


Figure 3.7: Best-Fit Model by Market Competition

The figure shows standardized structural coefficients of the best-fit model, categorised by LOW and HIGH market competition firms, along with descriptive and model statistics. Competition groups formed on basis of investor concentration (INV_Conc) and split at the 50th percentile. Information extracted from Table 3.5. All coefficients are significant at the 5% level or better.

Sampling Periods. To examine if regulatory reform (e.g. Regulation Fair Disclosure 2000, Sarbanes-Oxley Act 2002) and stock market turmoil (dot-com bubble 2000/02, financial crisis 2007/08) in the early 2000s affected the importance of idiosyncratic information for market participants, I split the sample into two periods (1993/99; 2000/10) and compare structure coefficients across them to test if there are significant structural differences between the 1990s and 2000s (all differences significant at the 1% level, see Table 3.7). Most structure coefficients for the two periods are widely similar in terms of their economic significance; however, the direct effect of *Asymmetry* on a firm's *CoE* is twice as strong in the 2000s (0.667, p-value: 0.000) than in the 1990s (0.372, p-value: 0.002), with the 2000s period explaining 76.1 percent in CoE variation, compared to only 39.6 percent for the 1990s (see Figure 3.8).

One explanation might be that regulatory change in the beginning of 2000 made private information more precious, causing more—not less—information asymmetry between market participants, resulting in firms' being expected to offer higher returns to investors.[55] While findings in Duarte et al. (2008) and Gomes et al. (2007) broadly support this argument, convincing evidence also exists to the contrary: for instance, Lee et al. (2014) show that regulatory reforms reduced security mispricing and Chen et al. (2010) document a decrease in cost of equity in the post-Reg-FD period. My methodology does not allow for further investigation of this increased *Asymmetry* effect and, thus, any strong conclusions in this regard would be premature. On a more general level, though, results indicate that—irrespective of market conditions and regulatory reform—firms' information quality remains of material importance to investors.

Descriptive Statistics		
	1993/99	2000/10
Age (yrs)	29.08	27.47
Listing (yrs)	13.00	15.35
Market Cap. (m$)	2,656.83	4,515.22
ROA (%)	2.75	1.35
Model Statistics		
CoE R^2	0.396	0.761
CFI	0.840	0.840
SRMR	0.090	0.090
Obs.	21,475	39,520

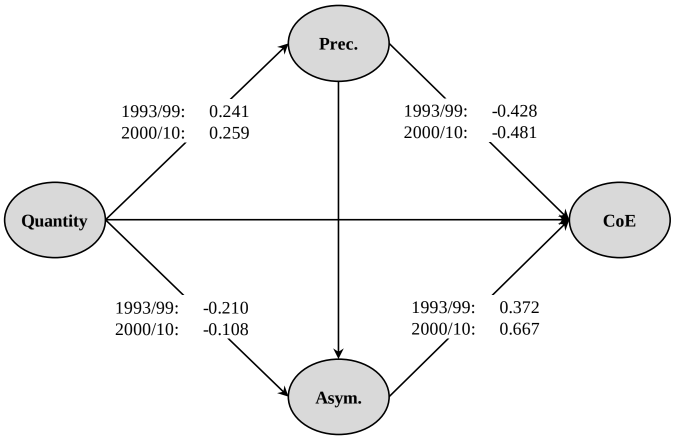


Figure 3.8: Best-Fit Model by Sampling Periods

The figure shows standardized structural coefficients of the best-fit model, split into sampling periods (1993/99; 2000/10), along with descriptive and model statistics. Information extracted from Table 3.5. All coefficients are significant at the 5% level or better.

To sum up, findings substantiate the proposition that firms with high (low) quality information environments enjoy relatively low (high) costs of equity. More specifically, results suggest that (1) CoE relevance of idiosyncratic information is negatively correlated with firm size, maturity and profitability, since younger, smaller and less profitable

55 To reduce information asymmetry (mainly between large institutional and small retail investors), the SEC enacted Regulation Fair Disclosure (FD) that requires companies to disclose material information to all investors at the same time. Let us assume that for fear of lawsuits over not releasing material information to all investors on time, companies not only disclose less overall information, but in particular starkly reduce disclosing private information to selected investors, just as FD intended. Under such regulation, it is likely that very few investors can maintain their access to informal (*alias* private) information, which leads to an enlarged portion of uninformed (*alias* publicly informed) investors. While under this new rule (public) information is now simultaneously disseminated among all investors, there is less information to be distributed in the first place—a concern raised by SEC commissioner Unger (2000) and investment professionals alike (AIMR, 2001). Having an increased fraction of publicly informed investors whose only source of information is now diminished results in greater information asymmetry and thus induces higher cost of equity (EO, 2004).

firms tend to benefit to a much greater extent from investing in the quality of their information environments than older, bigger and more profitable firms do; and (2) information risk—in the form of *Asymmetry* and *Precision* effects—tends to decrease with market competition as it facilities quicker revelation of private information in stock prices, which lowers firms' costs of equity.

3.4.3 Results of CoE Analyses

The proceeding paragraphs examine the extent to which the informational part of the best-fit SEM model explains variations in different CoE measures. This is an interesting inquiry, because if firm-specific information is indeed of varying relevance for VMB proxies compared to RFB proxies, then idiosyncratic information might offer an alternative explanation for why the former proxies are more valid measures of expected returns than the latter (Botosan and Plumlee (2005), Botosan et al. (2011), Lee et al. (2010, 2015)). Table 3.8 reports main results along with fit statistics for the different CoE measurement combinations. Based on these results, I provide several figures to ease interpretation of my findings.

3.4.3.1 Measurement Variations

I begin this analysis by separately estimating the best-fit SEM model for three CoE constructs: (1) CoE-RFB which is composed of the RFB proxies *rCAPM*, *rFF3*, *rFF4;* (2) CoE-FVIX which includes the VIX augmented RFB measures *rFVIX*, *rFVIX3*, *rFVIX4*; and (3) CoE-VMB which reflects the ICC measures *rPE*, *rPEG*, and *rAEGM*. The focus is on changes in the direct CoE effects (specified in H1-H3), given that all other structure coefficients (i.e., *Quantity* on *Precision*; *Quantity* on *Asymmetry*; *Precision* on *Asymmetry*) are unchanged under varying CoE measurement (see Table 3.8, Panel A).

Figure 3.9 shows results for the CoE-RFB model. The three factor loadings for *rCAPM* (0.455), *rFF3* (1.313) and *rFF4* (0.567) conform to expectations (positive and significantly non-zero) and show high internal consistency (Cronbach's alpha: 0.77). However, the direct CoE effects are either incompatible with predictions or insignificant. First, the positive association between *Precision* and *CoE* leads to the rejection of hypothesis H1 (0.046; p-value: 0.007). Second, the direct CoE effects stemming from *Asymmetry* and

Quantity are statistically and/or economically weak (*Asymmetry*: 0.016, p-value: 0.037; *Quantity*: -0.036; p-value: 0.112). As a result, the model explains a mere 0.3 percent of CoE-RFB variation and displays rather poor model fit-statistics. Findings for CoE-FVIX are similar to CoE-RFB (see Figure 3.10).

Figure 3.11 depicts factor loadings and structure coefficients for the CoE-VMB model. The construct loads significantly positive on all three ICC measures (*rPE*: 0.694; *rPEG*: 1.044; *rAEGM*: 0.320), with a Cronbach's alpha of 0.40. The direct effect between *Quantity* and *CoE* is inconsistent with H3, but negligible from an economic perspective (0.102, p-value: 0.000). More importantly, *Precision* (-0.353, p-value: 0.000) and *Asymmetry* (0.323, p-value: 0.000) both constitute an economically and statistically meaningful impact on CoE and are in line with hypotheses H1 and H2. This leads to a well-fitted model that explains 24.4 percent of firms' expected returns.

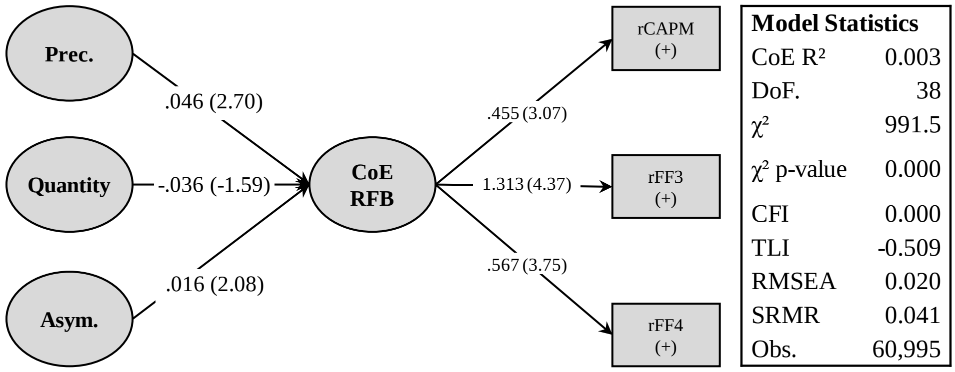


Figure 3.9: RFB CoE Construct

The figure shows factor loadings for the risk factor/based CoE construct and standardized structure coefficients for Precision, Quantity and Asymmetry. T-statistics based on standard errors clustered by firm reported in parentheses.

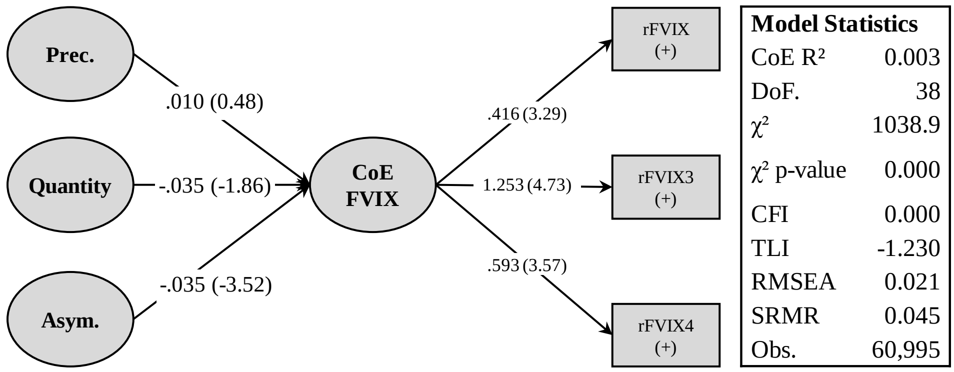


Figure 3.10: FVIX CoE Construct

The figure shows factor loadings for the VIX augmented risk factor/based CoE construct and standardized structure coefficients for Precision, Quantity and Asymmetry. T-statistics based on standard errors clustered by firm reported in parentheses.

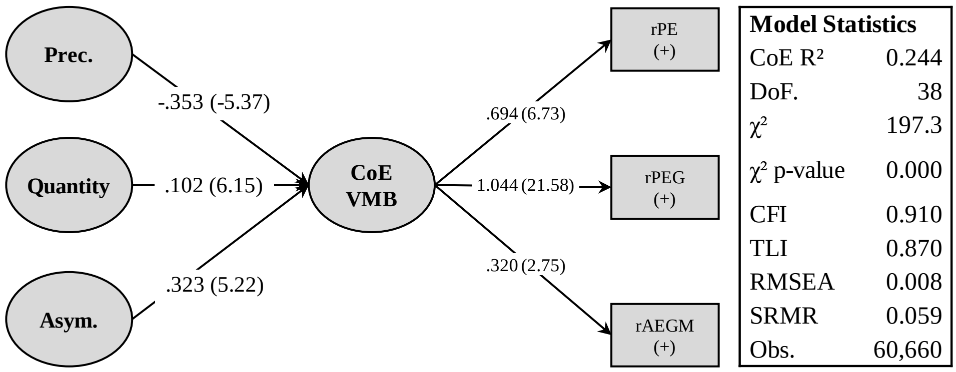


Figure 3.11: VMB CoE Construct

The figure shows factor loadings for the valuation model/based CoE construct and standardized structure coefficients for Precision, Quantity and Asymmetry. T-statistics based on standard errors clustered by firm reported in parentheses.

Table 3.8: CoE Measurement Variations

Panel A: CoE Measurement Variations & Structural Coefficients (standardized)

	Risk-Factor-Based (RFB)				**Risk Factor-Based + VIX (FVIX)**				**Valuation-Model-Based (VMB)**			
CoE Factor Loadings	RFB1	RFB2	RFB3	RFB4	FVIX1	FVIX2	FVIX3	FVIX4	VMB1	VMB2	VMB3	VMB4
rCAPM	0.455		x	x								
rFF3	1.313	0.757	x									
rFF4	0.567	0.984		x								
rFVIX					0.416		x	0.284				
rFVIX3					1.253	0.731	x					
rFVIX4					0.593	1.016		0.864†				
rPE									0.694	0.571		1.281†
rPEG									1.044	1.264	1.001	
rAEGM									0.320		0.321	0.064†

Direct Effects	RFB1	RFB2	RFB3	RFB4	FVIX1	FVIX2	FVIX3	FVIX4	VMB1	VMB2	VMB3	VMB4
Quant. → Prec.	0.283	0.279	-	-	0.283	0.276	-	0.275	0.276	0.275	0.277	0.276
Quant. → Asym.	-0.239	-0.239	-	-	-0.240	-0.240	-	-0.240	-0.240	-0.242	-0.241	-0.242
Prec. → Asym.	-0.120	-0.120	-	-	-0.119	-0.119	-	-0.119	-0.120	-0.117	-0.116	-0.109
Quant. → CoE	-0.036†	-0.100	-	-	-0.035*	-0.086	-	-0.098†	0.102	0.073	0.106	0.003†
Asym. → CoE	0.016	-0.031†	-	-	-0.035	-0.056	-	-0.059	0.323	0.209	0.342	0.295†
Prec. → CoE	0.046	0.086	-	-	0.010†	0.098	-	0.119†	-0.353	-0.315	-0.343	0.279†

Table continued on next page.

Table 3.8: CoE Measurement Variations(cont.)

Panel B: Explained Variance (R-Square)

	Risk-Factor-Based (RFB)				Risk Factor-Based + VIX (FVIX)				Valuation-Model-Based (VMB)			
Dependent Variable	RFB1	RFB2	RFB3	RFB4	FVIX1	FVIX2	FVIX3	FVIX4	VMB1	VMB2	VMB3	VMB4
CoE	0.003	0.013	-	-	0.002	0.015	-	0.020	0.244	0.151	0.249	0.136
Prec.	0.080	0.078	-	-	0.080	0.076	-	0.076	0.076	0.076	0.077	0.076
Asym.	0.088	0.087	-	-	0.088	0.087	-	0.087	0.088	0.087	0.088	0.085

Panel C: Fit Statistics

	Risk-Factor-Based (RFB)				Risk Factor-Based + VIX (FVIX)				Valuation-Model-Based (VMB)			
Statistic	RFB1	RFB2	RFB3	RFB4	FVIX1	FVIX2	FVIX3	FVIX4	VMB1	VMB2	VMB3	VMB4
DoF.	38	29	-	-	38	29	-	29	38	29	29	29
χ^2	991.5	886.8	-	-	1038.9	1348.6	-	1011.1	197.3	134.9	93.8	98.9
χ^2 p-value	0.000	0.000	-	-	0.000	0.000	-	0.000	0.000	0.000	0.000	0.000
CFI	0.000	0.000	-	-	0.000	0.000	-	0.000	0.910	0.974	0.943	0.960
TLI	-0.509	-1.443	-	-	-1.230	-3.864	-	-1.398	0.870	0.960	0.912	0.938
RMSEA	0.020	0.022	-	-	0.021	0.027	-	0.042	0.008	0.008	0.006	0.006
SRMR	0.041	0.037	-	-	0.045	0.039	-	0.042	0.059	0.071	0.036	0.058
Obs.	60,995	60,995	-	-	60,995	60,995	-	60,995	60,660	60,660	60,659	60,660

*† p > 0.10, * p > 0.05. All other coefficients are significant at the 5% level or better. All estimates standardised. For RFB3, RFB4 & FVIX3 model failed to converge.*

3.4.3.2 Measurement Combinations

To elaborate on the role of idiosyncratic information in the measurement of CoE, I simultaneously reflect RFB, FVIX and VMB proxies in the measurement model (Table 3.9). This allows checking for measurement error in the different proxies and controls robustness to my findings.

Figure 3.12 illustrates factor loadings and structure coefficients for the RFB-VMB CoE construct. As shown the VMB estimates load positively and highly significantly on the respective CoE construct (*rPE*: 0.728; *rPEG*: 0.996; *rAEGM*: 0.349), while loadings for the RFB proxies are negative and insignificant (*rCAPM:* -0.128; *rFF3:* -0.104; *rFF4*: -0.125). Amending the RFB estimates by an additional risk factor for market-wide volatility expectations (FVIX) leaves results unchanged (see Figure 3.13).[56] More importantly, structure coefficients and explained variance of CoE for both RFB-VMB and FVIX-VMB are almost identical to the results for the VMB-CoE construct stand-alone (see Figure 3.11). This is direct evidence that valuation model-based proxies are more likely to capture changes in firms' information environments than risk factor-based proxies do.

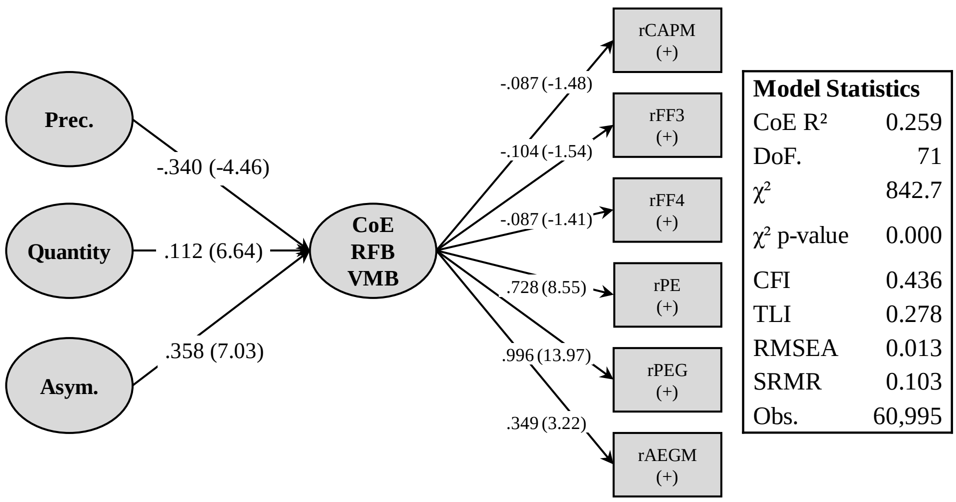


Figure 3.12: RFB-VMB CoE Construct

The figure shows factor loadings for the RFB-VMB CoE construct and standardized structure coefficients for Precision, Quantity and Asymmetry. T-statistics based on standard errors clustered by firm reported in parentheses.

[56] I do not report results for the RFB-FVIX combination given that results are alike to measuring CoE by RFB (Figure 3.9) and FVIX (Figure 3.10) stand-alone. This is due to due to high correlations of 0.93 (0.87) [0.94] between rCAPM and rFVIX (rFF3 and rFVIX3) [rFF4 and rFVIX4].

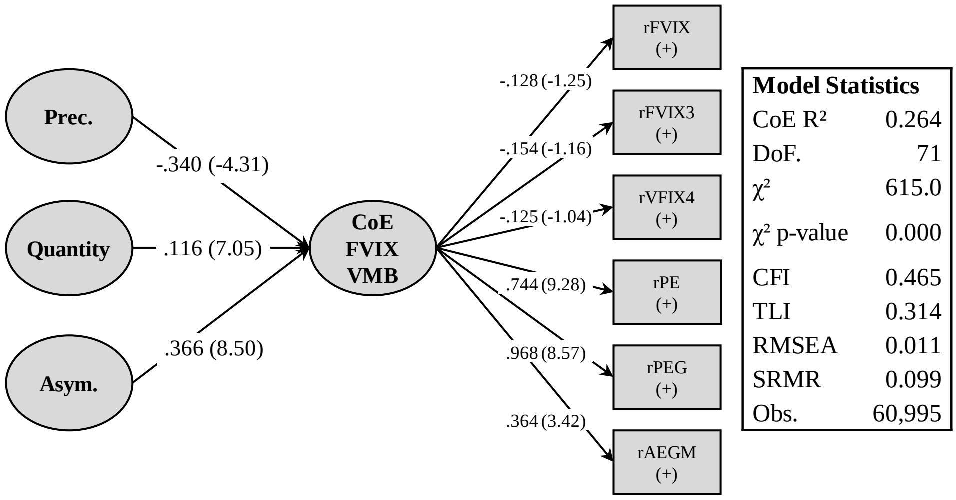


Figure 3.13: FVIX-VMB CoE Construct

The figure shows factor loadings for the FVIX and VMB CoE construct and standardized structure coefficients for Precision, Quantity and Asymmetry. T-statistics based on standard errors clustered by firm reported in parentheses.

The fact that RFB proxies reflect market-wide CoE expectations (as based on stock price equilibria and, in the case of FVIX, volatility outlooks), while VMB proxies capture investors' sentiment at the firm level (in the form of analyst forecasts) might clarify why firm-specific information is markedly reflected in VMB estimates, but not in RFB/FVIX ones: VMB estimates derive from analyst forecasts, with security analysts likely to pay greater attention to (or at least show greater sensitivity for) the quality of firms' information environments than average market participants, which explains their stronger association with information attributes vis-à-vis RFB estimates. Differently stated, the impact of idiosyncratic information on security analysts' expectations is much greater than on the market as a whole; whether this is good or bad depends on the extent to which the inclusion of firm-specific information adds to the construct validity of the various proxies. With the comparative literature showing that VMB outperforms RFB measures (Botosan and Plumlee (2005), Botosan et al. (2011), Lee et al. (2010, 2015)), the incorporation of idiosyncratic information in the measurement of CoE might increase the empirical soundness of those estimates.

Moreover, among the three RFB-VMB combinations (see Table 3.9), the one consisting of just *rFF4* and *rPEG* reflects the most parsimonious and best-fitted SEM specification (i.e., fit-statistics improve as the number of CoE proxies decreases). Interestingly, the SEM model explains about 50 percent more CoE variation in this 1 by 1 combination

(R^2: 0.413) than in the 2 by 2 (0.244) and 3 by 3 (0.259) ones. Consistent with the previous argument, the reason might be that among the VMB proxies, *rPEG* reflects the greatest and purest fraction of analysts' sentiment because it is exclusively estimated from one- and two-year ahead consensus earnings forecasts, while the others either only use one-year ahead forecasts (*rPE*) or make ad hoc assumptions unrelated to analysts' forecasts (e.g., perpetual growth in *rAEGM*).

Overall, findings show that firm-specific information is strongly reflected in VMB estimates, but not so in RFB/FVIX ones. A possible explanation is that VMB measures are calculated from both idiosyncratic and systematic information, while RFB/FVIX proxies derive from systematic information only.[57] Given prior evidence that VMB tends to outperform RFB proxies, I argue that the inclusion of idiosyncratic information in the measurement of CoE might help increase the construct validity of expected return proxies. A first promising step in this direction is taken by Ecker et al. (2006).

---

[57] It seems likely that the explanatory power of FVIX proxies would markedly increases if one amends RFB proxies by implied volatility estimates on the firm-level, instead of market-wide VIX data as in this study; however, due to data restrictions, we cannot carry out this analysis.

Table 3.9: Selected CoE Measurement Combinations

Panel A: CoE Measurement Variations & Structural Coefficients (standardized)

			RFB x VMB			FVIX x VMB		
CoE Factor Loadings			3x3	2x2	1x1	3x3	2x2	1x1
rCAPM			-0.087†	-0.049				
rFF3			-0.104†					
rFF4			-0.087†	-0.062	-0.070			
rFVIX						-0.128†	-0.081	
rFVIX3						-0.154†		
rFVIX4						-0.125†	-0.088*	-0.098
rPE			0.728			0.744		
rPEG			0.996	1.013	0.777	0.968	0.912	0.601
rAEGM			0.349	0.316		0.364	0.355	
Direct Effects			3x3	2x2	1x1	3x3	2x2	1x1
Quant.	→	Prec.	0.276	0.277	0.277	0.276	0.276	0.275
Quant.	→	Asym.	-0.240	-0.241	-0.241	-0.240	-0.242	-0.241
Prec.	→	Asym.	-0.120	-0.116	-0.116	-0.121	-0.115	-0.115
Quant.	→	CoE	0.112	0.106	0.139	0.116	0.119	0.181
Asym.	→	CoE	0.358	0.342	0.437	0.366	0.368	0.555
Prec.	→	CoE	-0.340	-0.337	-0.447	-0.340	-0.381	-0.581

Panel B: Explained Variance (R-Square)

	RFB x VMB			FVIX x VMB		
Dependent Variable	3x3	2x2	1x1	3x3	2x2	1x1
CoE	0.259	0.244	0.413	0.264	0.297	0.683
Prec.	0.076	0.077	0.077	0.076	0.076	0.076
Asym.	0.088	0.087	0.087	0.088	0.087	0.087

Panel C: Fit Statistics

	RFB x VMB			FVIX x VMB		
Statistic	3x3	2x2	1x1	3x3	2x2	1x1
DoF.	71	48	29	71	48	29
χ^2	842.7	103.6	147.3	615.0	114.5	110.5
χ^2 p-value	0.000	0.000	0.000	0.000	0.000	0.000
CFI	0.436	0.908	0.959	0.465	0.901	0.966
TLI	0.278	0.874	0.936	0.314	0.863	0.947
RMSEA	0.013	0.004	0.008	0.011	0.005	0.007
SRMR	0.103	0.048	0.038	0.099	0.047	0.037
Obs.	60,995	60,995	60,995	60,995	60,995	60,995

*† $p > 0.10$, * $p > 0.05$. All other coefficients are significant at the 5% level or better. All estimates standardised. For FVIX x VMB (3x3) model failed to converge.*

3.5 Summary and Conclusions

The empirical literature examining the repercussions of idiosyncratic information on firms' CoE is voluminous, and the conclusions reached vary widely depending on the proxies used to estimate information quantity, asymmetry, precision and CoE. By applying a SEM approach—which allows for contemporaneous analysis of the most commonly used empirical measures—I bring new evidence to the discussion about the pricing of information risk (e.g., Core et al. (2008), Mohanram and Rajgopal (2009), Shevlin (2013)) and the construct validity of different CoE estimates (e.g., Botosan and Plumlee (2005), Easton and Monahan (2016)).

In particular, this study complements BEOS (2012) in that it empirically validates and simultaneously tests for the analytical predictions of EO (2004) and LLV (2012). However, it differs from their path analyses in that my SEM approach is not bound to using only one direct proxy per construct, but has the benefit to concurrently rely on several indicators for *Quantity*, *Asymmetry*, *Precision* and *CoE*. This enhances measurement reliability of these constructs (as only reflecting shared variance among the indicators) and thus the soundness of this study's empirical conclusions.

Initial results from estimating the conceptual SEM model corroborate the proposition that firms with high (low) quality information environments tend to enjoy relatively low (high) costs of equity. More specifically, results show that firms' information environments are dominated by the direct impact of *Precision* (-0.447) and *Asymmetry* (0.437) on CoE (structure coefficients in brackets). *Quantity*, on the other hand, appears to be of limited importance (0.139), but it serves as the precursor of *Precision* (0.277) and *Asymmetry* (-0.242); that is, only as the amount of available information about a firm increases (e.g., in the form of mandatory and voluntary disclosure) can investors begin to evaluate the accuracy of this information and informational disadvantages between informed and uninformed investors can be resolved. Relaxing the assumption that information precision is mediated by asymmetry (as maintained in BEOS, 2012), and given that both attributes show an equally strong impact on CoE, I conclude that if firms must decide to either provide more accurate information to investors (increasing *Precision*) or distribute private information more equally between them (decreasing *Asymmetry*), they should dedicate scarce resources to those corporate activities that show the most room for improvement (e.g., investor relations, reporting quality/frequency).

In analysing the relative importance of idiosyncratic information, the following insights are provided. First, it is shown that the younger and smaller firms are the greater their CoE benefits from actively investing in high quality information environments. This can be inferred from decreasing levels of explanatory power of the model as one moves from AMEX (CoE R^2: 0.603) to NASDAQ (0.339) to NYSE (0.182) stocks, with the latter being much larger, older and more profitable than the former two. Second, the impact of information risk on firms' CoE decreases as market competition (in the form of less investor concentration) increases; that is, *Asymmetry* effects are almost four times weaker for the high competition group (0.058) than for the low competition group (0.209), and *Precision* effects are about 1.5 times weaker (high: -0.519; low: -0.328). Taken together, results indicate that CoE effects stemming from *Asymmetry* are approximately twice as sensitive to changes in firm size, maturity and market competition than *Precision* (i.e., decaying at a much faster rate), which qualifies my conclusion as follows: if less established firms face a trade-off between providing better quality information to investors and aiming to reduce informational disadvantages between investor groups, they should choose the former over the latter, given the greater persistence (i.e., CoE benefits) of *Precision* effects.

The second part of this study examines the explanatory power of firm-specific information for different expected return proxies. Findings suggest that valuation model-based CoE measures impound significantly more idiosyncratic information than risk factor-based estimates; that is, informational differences between firms explain substantial variation in the three VMB estimates (CoE R^2: 0.244), but none in the RFB (0.003) and FVIX (0.003) proxies. Combining this finding with results from the performance literature (in which it is shown that VMB outperform RFB proxies) might partially explain the difference in construct validity between the two groups: VMB proxies capture investors' sentiment (in the form of analysts' forecasts) on the firm-level, while RFB proxies (based on stock price equilibria) reflect market-wide CoE expectations, which nominates firm-specific information as the missing explanatory link between the two. That is, the accuracy and accessibility of firm-specific information directly affects analysts' forecasts, but certainly less so—if at all—the average market participant. On the one hand, this explains the much stronger association between information attributes and VMB proxies vis-à-vis RFB estimates; on the other hand, it supports the notion that the incorporation of idiosyncratic information in the measurement of RFB proxies might improve the empirical soundness of those estimates.

Overall, this study contributes for a large sample of 7,091 firms confirmatory evidence that firm-specific information constitutes a significant statistically and economically impact on firms' cost of equity. However, results also show that the strength of this impact tends to be conditional on firm size, maturity, profitability as well as market competition, and only becomes pronounced in analyst-based ICC estimates, but not in market-based return proxies. Therefore, quantifying this impact via an information risk factor, which contemporaneously captures firms' asymmetry and precision characteristics, could be an interesting path for future asset pricing research.

3.6 Appendices

Appendix 3.1: SEM Approach-Based Investigation of the Link between Idiosyncratic Information and Cost of Equity

This paper examines the link between idiosyncratic information and cost of equity by means of a structural equation modelling (SEM) approach, which appears to be most appropriate given the peculiarities of this study. On the one hand, the focus lies on information constructs which are unlikely to be directly measurable, and on the other hand, it is intended to provide an integrated view on the extent to which informational differences across firms can (directly and indirectly) impact investors CoE expectations. One key strength of SEM is that it simultaneously extends regression and factor analysis in that it allows for dependent and independent latent constructs in the regression analysis while incorporating the notion of factor analysis, which is usually limited to the investigation of correlation structures between observable (non-latent) variables. In that sense, SEM merges multiple-regression and confirmatory factor analysis into one analytical step (Savalei and Bentler, 2010).

SEM models consists of a structural and a measurement model. The structural model usually derives from existing theory and depicts the relation between the latent constructs. The measurement model infers the latent constructs by using a range of observable indicators for them. Each latent construct needs to be specified by at least two indicator variables (alias manifest or reference variables). The validation of the measurement model, i.e., verifying that the indicator variables measure the latent construct reliably, is a prerequisite before the full SEM model can be tested. This first step is referred to as "confirmatory factory analysis" (CFA) at which each indicator variable is treated as to have two causes: a single factor that measures the latent construct and all other unique sources of influence (omitted causes) represented by the error term. The shared variance (commonality) among the indicator variables reflects the degree to which the indicators measure the underlying factor/latent construct Eventually, a good SEM model is one which shows greatest resemblance between observed covariances in the data and predicted covariances by the model (Kline, 2011).

Model fit is assessed on how well a particular SEM specification resembles observed variance-covariance in the sample data with predicted variance-covariance by the model. Given the inapplicability of the chi-square statistic in large sample studies like this one, I evaluate model fit based on the following approximate fit indexes: standardised root mean square residual (SRMR), root mean square error of approximation (RMSEA), Tucker-Lewis-Index (TLI) and comparative fit index (CFI).[58] These indexes offer a continuous measure for how well the model corresponds with observed data and are generally scaled as "goodness-of-fit" measures (exception RMSEA), where higher values indicate better model-fit.

The SRMR captures misspecification of the structural part of the SEM, while TLI, CFI and RMSEA examine measurement model misspecifications. In line with Hu and Bentler's (1999) recommendation, I appraise model fit according to the following index levels; (1) TLI $\geq$ 0.95 and SRMR $\leq$ 0.09; (2) CFI $\geq$ 0.95 and SRMR $\leq$ 0.09; (3) RMSEA $\leq$ 0.06 and SRMR $\leq$ 0.09; but acknowledge that these levels shall not be mistaken as golden rules of model fit (Fan and Sivo, 2005). What is more, Kline (2011, p. 171) suggests to scrutinise correlation residuals (calculated as the difference between model-implied and observed-sample correlations) in order to detect those sample correlations which are not well explained by the overall model. As a general rule, absolute differences greater than 0.10 are being regarded as problematic in the SEM literature.

Appendix 3.2: Description of Easley and O'Hara (2004)

In a multi-asset, partially revealing rational expectations equilibrium model, Easley and O'Hara (EO, 2004) demonstrate that a firm's CoE is decreasing in information precision (i.e., decreasing in accuracy of the available information about the future value of the

[58] The model chi-square statistic is the most commonly used model test statistic in SEM research. It is structured as a "badness-of-fit" measure in that higher values indicate worse model fit: significant results (say, $p < 0.05$) denote overall model misspecification. However, unless perfect model fit is attained– which is unlikely in any real world application – the model chi-square statistics increases with sample size. Thus, in very large samples (e.g. N = 5,000, Kline, 2011, p. 201) even minor model-data discrepancies can lead to test statics rejecting an otherwise valid model (Fan et al., 1999). With number of observations well exceeding this threshold (N = 60,995 for most of my models), the model chi square statistic is an invalid measure of model fit in this study.

firm) and increasing in information asymmetry (i.e., increasing in the fraction of uninformed investors and private signals about the future value of the firm). The EO model considers two type of risk-averse investors; (1) the informed, who have access to both public and private information about the future value of a risky asset; and (2) the uninformed, who only have access to the public information set, but can partially infer some of the private information about the future value through price observations. Equation (3.1) duplicates the EO model along with a definition of the main parameters.

$$E[v_k - p_k] = \frac{\delta \bar{x}_k}{\rho_k + (1 - \alpha_k) I_k \gamma_k + \mu_k \alpha_k I_k \gamma_k + (1 - \mu_k) \rho_{\theta k}} \tag{3.1}$$

where

$E[v_k - p_k]$	*expected rate of return as the difference between future value of stock k (v_k) and buy price of stock k (p_k)*
$\bar{x}_k$	*the average per capita supply of stock k in number of shares*
δ	*coefficient of risk aversion*
μ_k	*fraction of privately informed investors;* $(1 - \mu_k)$ *uninformed investors*
ρ_k	*precision of prior information*
I_k	*signals about future value* v_k *of stock k*
γ_k	*precision of signals* I_k
α_k	*fraction of private signals in* I_k*;* $(1 - \alpha_k)$ *public signals*
$\rho_{\theta k}$	*precision of information revealed through trading*

Changing values in the numerator and/or the denominator have a direct impact on the expected rate of return of the risky asset (viz. stock *k*). Focussing on the numerator first (and holding parameters in the denominator constant) it is apparent that the CoE of share *k* increases with investors risk aversion (δ). Correspondingly, if investors must hold a greater (average) number of risky shares *k* in their portfolios ($\bar{x}_k$), it follows that the risk associated with share *k* is less widely spread among investors which induces greater risk for every single investor; hence, CoE is increasing in the average per capita supply of the stock.

The major contribution of the model, however, stems from predictions obtained by altering parameter values in the denominator (while assuming parameters in the numerator remain unchanged). The model predicts that the CoE is a decreasing function of stock's k information structure, given by

$$\rho_k + (1 - \alpha_k)I_k\gamma_k + \mu_k\alpha_k I_k\gamma_k + (1 - \mu_k)\rho_{\theta k} \tag{3.2}$$

which can be rewritten as and is tantamount to

$$\mu_k(\rho_k + I_k\gamma_k) + (1 - \mu_k)(\rho_k + (1 - \alpha_k)I_k\gamma_k + \rho_{\theta k}) \tag{3.3}$$

Eq. (3.2) reveals the main subsets of information that constitute a firm's information environment: "prior information with precision ρ_k, public information with precision $(1 - \alpha_k)I_k\gamma_k$, private information with precision $\alpha_k I_k\gamma_k$, public plus private information with precision $I_k\gamma_k$ and private information partially revealed through price with precision $\rho_{\theta k}$" (Botosan and Plumlee, 2013, p. 1047). As it can be seen from Eq. (3.2), an increase in the precision parameter (γ_k) generates a positive impact on the entire information structure, leading to an increased level of information precision in both private and public information and eventually lowers the CoE.

The impact of asymmetric information on the CoE is best illustrated with reference to Eq. (3.3). On the one hand, asymmetric information is influenced by how widely the information is disseminated among investors (dissemination: μ_k) and on the other hand by the extent to which information is truly private and, therefore, only accessible by the informed investors (composition: α_k). It is apparent from Eq. (3.3) that if the fraction of privately informed investors (μ_k) increases, more weight is placed on the richer information set $I_k\gamma_k$ which increases the "weighted average assessed precision of information in the market" (ibid.) and, therefore, decreases the CoE. The impact of the composition of information (α_k) on CoE, however, depends on the assumptions made about I_k.

In the EO model the information content of I_k (i.e., the sheer quantity of signals about the future value of share k) is assumed to be constant (EO, 2004, p. 1558).[59] Varying α_k then simply determines the fraction of I_k which is private and public knowledge; that is, variations in α_k determine the degree to which public information is substituted by private information (vice versa). As elucidated in Eq. (3.3), informed investors are unaffected by changing levels of α_k as they observe all available information about stock k. However, increasing levels in α_k decrease both the quantity and precision of available information on which the uninformed can base their expectations about the future value of share k. This informational disadvantage increases information asymmetry of stock *k* which translates into higher expected rate of returns.

In a direct critique of EO (2004), Hughes et al. (2007) present a model in which the risk from information asymmetry is fully diversifiable and such is no longer a priced risk factor on its very own, but impacts the CoE through its effect on systematic risk factor loadings—foremost market beta. Information asymmetry is then sill influential, but now mediated by systematic risk factors. Such a model is certainly in greater alignment with neoclassic asset pricing theory, which states that idiosyncratic risk is fully diversifiable in large economies and, therefore, it is only systematic risk that is priced in the markets. However, as long as beta is measured with error (i.e., a well-specified, forward-looking beta which subsumes the risk from information asymmetry is not attainable), than information asymmetry might still appear as a separate risk factor in empirical work (Lambert et al., 2007). Moreover, Hughes et al. (2007, p. 723) themselves acknowledge that their theory is “not inconsistent with studies that presume an existence of a systematic ‘information risk’ factor.”

[59] This assumption is relaxed in Lambert et al, 2012.

Appendix 3.3: Description of Lambert, Leuz and Verrecchia (2012)

Lambert, Leuz and Verrecchia (LLV, 2012) develop a single-asset, partially revealing rational expectations equilibrium model which is constituted of risk averse informed and uninformed investors.[60] Although their major focus lies on the impact of asymmetric information on the cost of capital under varying levels of market competition, their model is general enough to allow for a comprehensive investigation of the impact of all information attributes on firms' CoE. The main argument of the LLV model is that in perfectly liquid markets where all investor types act as price takers—that is, perfect market competition in the form of "horizontal demand curves for stocks" (Shleifer, 1986, p. 579) exist—asymmetric information has no impact on the CoE *over and above* its impact on average precision about the future value of the risky asset. However, if markets are imperfect there is a separate cost of equity effect induced by asymmetric information, even when the average precision in the market remains unchanged. The LLV model is shown in Eq. (3.4) along with a definition of the main parameters.

$$E[\tilde{V}] - E[\tilde{P}] = \left[\frac{(1 + r_I \Pi_I \lambda)^{-1} N r_I \Pi_I + M r_U \Pi_U}{N r_I + M r_U}\right]^{-1} E\left[\frac{\tilde{Z}}{N r_I + M r_U}\right] \tag{3.4}$$

where

$E[\tilde{V}] - E[\tilde{P}]$	*expected rate of return as the difference between future cash flow of the risky asset ($\tilde{V}$) and price of shares in the risky asset ($\tilde{P}$)*
N, M	*number of informed and uninformed investors, respectively*
r_I, r_U	*risk tolerances of informed and uninformed investors, respectively*
$\Pi_I = \Pi_V + \Pi_\epsilon$	*precision of informed investors' beliefs about the risky asset's cash flow*
$\Pi_U = \Pi_V + \Pi_\delta$	*precision of uninformed investors' beliefs about the asset's cash flow*
λ	*an informed investor's coefficient of illiquidity*
$\tilde{Z}$	*supply of shares in the risky asset*

The LLV model measures competition as "the extent to which an informed investor anticipates that his demand order will move price" (ibid., p. 13) and is denoted by the

[60] Lambert and Verrecchia (2015) offer closed-form solutions to the equilibria conditions in the LLV (2012) model, however, leaving the analytical conclusions unaffected.

coefficient λ. The genuine assumption is made that if either the number of informed or uninformed investors becomes very large, market liquidity increases up to a point where it renders both investor types into price takers. This mechanism establishes a perfect market in which λ approaches zero. Furthermore, the precision parameter Π_I subsumes prior (Π_V) and private (Π_ϵ) information available to informed investors. The precision parameter Π_U captures prior information (Π_V) and information which uninformed can infer from price realisations (Π_δ).[61]

Perfect Market Competition. Assuming that markets are perfectly competitive (i.e. $\lambda = 0$), then the LLV model reduces to the form shown in Eq. (3.5) and allows for a direct comparison with EO (2004) model which also maintains the assumption of perfect market competition.

$$E[\tilde{V}] - E[\tilde{P}] = \left[\frac{Nr_I\Pi_I + Mr_U\Pi_U}{Nr_I + Mr_U}\right]^{-1} E\left[\frac{\tilde{Z}}{Nr_I + Mr_U}\right] \tag{3.5}$$

Holding the expectation about the supply of the risky asset constant, it follows from Eq. (3.5) that CoE is now a decreasing function of the weighted average of the informed and uninformed investors' information precision about the future cash flow of the risky asset; hence, the first term in Eq. (3.5) is simply referred to as average precision and denoted as:

$$\Pi_{avg} = \frac{Nr_I\Pi_I + Mr_U\Pi_U}{Nr_I + Mr_U} \tag{3.6}$$

The predictions from Eq. (3.5) are similar to these in EO (2004) with regards to information precision (γ_k), but are subtler with respect to information asymmetry which arises

[61] Public information is not explicitly considered in the LLV model which is insofar not much of an issue, as the main purpose is to demonstrate that the impact of asymmetric information on the CoE depends on the level of market competition. Asymmetric information stems from the fact that some investors have access to private information, while others do not. Therefore, the model would reach the same conclusions even when public information is incorporated into the model. The only difference would be that the posterior precision of each investor type (i.e. Π_I & Π_U) is increased by the "precision of the error term of the public signal" (LLV, 2012, p.7, footnote 6).

from the dissemination (μ_k) and composition (a_k) of information. The first prediction of the LLV model is that an increase in prior information (Π_V) induces an increase in precision of both the informed and uninformed investors through $\Pi_I = \Pi_V + \Pi_\epsilon$ and $\Pi_U = \Pi_V + \Pi_\delta$, respectively. This increases the average precision in the market which lowers the CoE of the firm.[62]

The second prediction—which constitutes the most crucial point of the LLV model—is that in perfectly competitive markets information asymmetry "has no impact on the cost of capital *after controlling* for any impact it might have on average precision" (ibid., footnote 12, emphasis added). Changing degrees of information asymmetry always operate through changes in average precision first. Only if these changes increase (decrease) the average precision about the future payoffs of the risky asset, then the CoE decreases (increases), irrespective of whether this change stems from more or less asymmetric information. The following two examples illustrate as to why information asymmetry and CoE do not always move in concert with each other.

First, assume that the composition of information is hold constant, but more investors become informed (increasing *N*). Then more weight is placed on the more accurate information set Π_I which increases average precision (Π_{avg}) and decreases CoE, while also *decreasing information asymmetry* (i.e. private information is now more widely disseminated among investors).[63] Second, assume that the fraction of privately informed investors remains unchanged, while "the amount of private information" (ibid., p.17) grows (increasing Π_ϵ). Such a shift not only increases the information set of the privately informed, but also the information set of the uniformed as they now can infer more information from price due to the arrival of a new private signal in the market (increasing Π_δ). Hence, the average precision in the market increases and the CoE decreases, despite a

62 Note, the same prediction is obtained if one increases the overall precision (γ_k) of available information in the EO model.

63 Note, that the same prediction is obtained from the EO model if one increases the fraction of privately informed investors (μ_k).

simultaneous *increase in information asymmetry* (i.e., more private information is available to the informed investors).[64] The fundamental insight from above is that in markets with perfect competition the "communication of more information to more investors, not the reduction of information asymmetry *per se*, lowers the cost of capital" (Lambert et al., 2012, p. 18). Stated differently, a decrease in information asymmetry can either increase or decrease the average precision about the future payoffs of the risky asset; if the latter happens the CoE increases.

Imperfect Market Competition. If markets are imperfect (i.e. $\lambda \neq 0$) than the average precision is not just the simple weighted average of the precision of the two investor types (as under perfect competition), but is also affected by the degree of market illiquidity as shown in Eq. (3.7).

$$\Pi_{avg} = \frac{(1 + r_I \Pi_I \lambda)^{-1} \, N r_I \Pi_I + \, M r_U \Pi_U}{N r_I + \, M r_U} \tag{3.7}$$

While mechanically obvious that the average precision decreases as illiquidity (λ) increases, the notion behind needs clarification. In imperfect markets the demand of the informed depends not only on their risk tolerance and information about the future payoffs of the risky asset, but also on the extent to which their demand impacts prices. As illiquidity in the markets increases, the informed are less aggressive in trading the stock as each trade has now a much greater price impact than under perfect market competition; that is, demand curves are no longer horizontal but downward sloping which means that price curves are now being upward sloping in demand (Shleifer, 1986). In such market settings, the informed are forced to behave much more strategically when to trade and how much to trade (Kyle, 1985). Therefore, less of their private information—which could be revealed if their demand would not impact price—is now reflected in stock

64 Note, that the opposite would be predicted in the EO model with regard to an increase in information asymmetry due to an increase in the composition of information (α_k). The differentiating feature between the two models is, that in the LLV model the information set is assumed to be variable; i.e. new private information also lead to an increase in the amount of information, while in the EO model the amount of information (I_k) is fixed; i.e. any changes in the composition of the information set is a simple substitution of private for public signals and *vice versa* (Botosan and Plumlee, 2013). However, both models reach the same conclusion if the fixed quantity assumption of the EO model is relaxed.

prices. This decreases the information precision of the uniformed investors—because part of their precision depends on the informational content of price realisations—which lowers total average precision and increases CoE.

To summarise, in imperfect market settings more private information (which also leads to greater information asymmetry) does not lower the CoE to the extent it could under perfect competition and this effect intensifies with increasing illiquidity. In other words, as illiquidity increases, the average precision which is assessed and available in the market is much lower than it could be if markets were perfectly liquid.[65]

Appendix 3.4: Empirical Measures of Information Quantity

The following three indicators are used to infer information quantity: period of listing (*Listing*); firm age (*Age)* and relative media coverage (*RMC*). All of them are assumed to be positively associated with information quantity and estimated as described below.

Period of Listing & Firm Age. Period of listing (firm age) is the number of years since a firm's initial public offering (incorporation) and is calculated as the difference between the fiscal year end the firm-year pertains to and the year of the IPO (incorporation) with incorporation dates being obtained from Osiris. IPO dates are provided by SDC Platinum, Osiris and CRSP, however, with varying degrees of firm coverage in each database. SDC Platinum is the industry standard for listing information; hence, I combine IPO year data according to the following hierarchy: 1st SDC Platinum, 2nd Osiris, 3rd CRSP; that is, in case no IPO year information is available on SDC Platinum, data provided by Osiris is used. If Osiris data is missing as well, CRSP stock header information is used instead. All observations for which the year of incorporation is lager then the IPO year are excluded from the sample.

65 To prove the reasoning above, one can specify the model such that an increase in information asymmetry leaves average precision unchanged, but at the same time increases CoE—a result which cannot be obtained in perfect markets (i.e. $\lambda = 0$). This implies that asymmetry has now an impact on CoE over and above its impact on average precision which can only stem from an increase in illiquidity (see LLV, 2012, numerical example in the appendix). This direct impact of information asymmetry on CoE increases with market illiquidity.

RMC Index. The RMC index is positive (negative) for firms that enjoy more (less) media coverage than the expected average firm in the same industry and of similar size. In estimating the RMC index, the methodology for the relative quantity index described in Beretta and Bozzolan (2004, 2008) and initially proposed in Beattie et al. (2002, 2004) is followed. The idea behind the RMC index is that variations in media coverage which are unexplained by industry-membership and firm size (i.e., the residuals) are a good proxy for firm prominence which, in turn, tends to be positively associated with higher information quantity (Kross and Schroeder, 1989). I proxy for media coverage of firm *i* as the number of news entries per firm-year on Factiva and use four-digit SIC codes to assign each firm into one of the 48 industries suggested by Fama and French (1997). I run the panel regression in Eq. (3.8)—which controls for time and industry effects—and obtain the RMC index per firm-year as shown in Eq. (3.9) below.

$$\widehat{MC}_{i,t} = \beta_0 + \sum_{j=1}^{T-1} \tau_j T_j + \sum_{j=1}^{k-1} \beta_j IND_j + \beta_k LNSIZE_{i,t} \tag{3.8}$$

$$RMC_{i,t} = MC_{i,t} - \widehat{MC}_{i,t} \tag{3.9}$$

where $\widehat{MC}_{i,t}$ ($MC_{i,t}$) is expected (observed) media coverage for firm *i* at time *t;* T_j and IND_j are time and industry dummies, respectively, $LNSIZE_{i,t}$ is the natural logarithm of sales (Compustat) and $RMC_{i,t}$ is the relative media coverage index for firm *i* at time *t.*

Appendix 3.5: Empirical Measures of Information Precision

Information precision is measured by two earnings quality indicators (viz. *accrual quality* and *earnings value relevance*) and one indicator for the precision of analyst forecasts. All indicators are assumed to be positively associated with information precision and their measurement is described below.

Earnings Quality. In this study, one accounting-based proxy *accrual quality* (*AQ*) and one market-based indicator *earnings value relevance (VR)* is used to proxy for firms' earnings quality.

Accrual quality metric. The most prevalent accrual quality metric in extant work is the one developed by Dechow and Dichev (2002) which is shown in Eq. (3.10). The idea behind this model is that accrual quality is higher for firms for which working capital accruals map better into previous, current and future operating cash flow realisations. A better mapping process is seen as indicative of low levels of abnormal accruals which is reflected by low regression residuals. The standard deviation of the residuals of the model are regarded as a "firm-level measure of accrual quality, where higher standard deviation denotes lower quality" (Dechow and Dichev, 2002, p. 40).

$$\Delta WC_t = \beta_0 + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \varepsilon_t \tag{3.10}$$

where ΔWC_t denotes change in working capital as a proxy for accruals, CFO_t is cash flow from operations in the current period. In this study, accrual quality (AQ) is based on the McNichols (2002) modification of the original Dechow-Dichev model as shown in equation (3.11) and measured as described in FLOS (2005, pp. 302-303).

$$\begin{aligned} TCA_{i,t} = \beta_{0,i} + \beta_{1,i} CFO_{i,t-1} + \beta_{2,i} CFO_{i,t} + \beta_{3,i} CFO_{i,t+1} \\ + \beta_{4,i} \Delta Sales_{i,t} + \beta_{5,i} PPE_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{3.11}$$

where $TCA_{i,t} = \Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STDEBT_{i,t}$ = total current accruals of firm *i* in year *t;* $CFO_{i,t} = NIBE_{i,t} - TA_{i,t}$ = cash flow from operations in current period;[66] $NIBE_{i,t}$ = net income before extraordinary items; $TA_{i,t} = \Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STDEBT_{i,t} - DEPN_{i,t}$ = total accruals; $\Delta CA_{i,t}$ = change in current assets between year *t-1* and *t*; $\Delta CL_{i,t}$ = change in current liabilities; $\Delta Cash_{i,t}$ = change in cash; $\Delta STDEBT_{i,t}$ = change in current/short-term debt; $DEPN_{i,t}$ = depreciation and amortization expense; $\Delta Sales_{i,t}$ = change in revenues and $PPE_{i,t}$= gross value of property, plant and equipment. All variables are deflated by firm *i'*s average total assets in year *t* and *t-1* ($Assets_{i,t}$) and winsorized at the 1st and 99th percentile in each year.

---

66 Total accruals are calculated from information provided in the balance sheet and income statement (i.e. indirect approach followed). As an alternative one could follows the direct approach favoured by Hribar and Collins (2002); however results are chiefly unaffected if data from cash flow statement are used (FLOS, 2005, footnote 3).

Annual cross-sectional estimations of Eq. (3.11) for each of the 48 Fama-French industries (1997) with at least 20 firms in year *t*, yield firm- and year-specific residuals which form the basis for the accrual metric. First, the standard deviations for each firms' residuals ($\varepsilon_{i,t}$) are calculated over years *t-4* through *t*, with large standard deviations indicating poor accruals quality and implying a negative relation with precision.[67] Second, I multiply the standard deviations by negative one in order to obtain the final accrual metric (*AQ*). This latter step differs from previous studies, but ensures a more intuitive interpretation between precision and accrual quality: the lower (higher) the information precision of a firm, the lower (higher) its accrual quality, the more (less) negative its standard deviation and *AQ* metric, respectively. That is, the *AQ* metric is now conjectured to be positively associated with precision, in contrast to an assumed negative association before the recoding.[68]

Earnings value relevance. An accounting amount—such as earnings—is regarded as value relevant if it is associated with stock prices in a predictable manner; that is, it bridges the gap between financial standard setters and equity markets (Barth et al., 2001, p. 79).[69] In this study, earnings value relevance (*VR*) is measured as the degree to which both a firm's earnings and change in earnings explain its stock returns, where greater explanatory power indicates more transparent and value relevant earnings (Barth et al. (2013), FLOS (2004)).

One way to operationalise *VR* is to measure the explained variability (adj. R^2) from regressing stock returns on price-deflated earnings and change of earnings for each firm

[67] Note that if a firm has consistently large residuals, so that the standard deviation of those residuals is small, that firm has relatively good accruals quality because there is little uncertainty about its accruals. For such a firm, the accruals map poorly into cash flows, but this is a predictable phenomenon, and should not be a reason for priced uncertainty (FLOS, 2005, p. 303).

[68] An alternative approach to estimate the AQ metric is to use firm-specific time-series regressions of annual data rather than annual industry cross-sections (see FLOS, 2004, pp. 979-980). However, this methodology requires an even longer time-series of firm-specific accounting information to be fully estimated. FLOS (2004) uses ten-years of firm data to estimate Eq. (3.11) in firm-specific time-series regressions, compared to seven years of firm data following the approach described above. This requirement would further reduce sample size, intensifies survivorship bias and thus excludes younger and smaller firms from analyses. However, in particular less mature firms constitute an interesting sample in which one would expect to observe greatest variety in accrual quality; thus, the industry approach taken in this study is to be preferred over this alternative time-series approach.

[69] See also Holthausen and Watts (2001) for an extended discussion of the value relevance literature.

over rolling ten-year windows as in FLOS (2004) or eight-year windows as in Bushman et al. (2004), with larger values of *VR* (i.e. larger adj. R^2) indicating more value relevant earnings.

$$RET_{i,t} = \beta_{0,i} + \frac{\beta_{1,i} NIBE_{i,t}}{MV_{i,t-1}} + \frac{\beta_{1,i} \Delta NIBE_{i,t}}{MV_{i,t-1}} + \varepsilon_{i,t} \tag{3.12}$$

where $RET_{i,t}$ = firm *i*'s 15-month continuously compounded return ending three month after the end of fiscal year *t*; $NIBE_{i,t}$ = net income before extraordinary items; $\Delta NIBE_{i,t}$ = change in net income before extraordinary items from year *t-1* to *t* and $MV_{i,t-1}$ = market value of firm *i* at the end of year *t-1* which is fiscal year closing stock price times common shares outstanding.

However, determining *VR* by means of firm-specific time-series regressions comes at the cost of a starkly diminished sample (see discussion of alternative measurement approach for accrual quality in footnote 68). Thus, I operationalise *VR* analogous the measurement of the accrual quality metric; that is, Eq. (3.12) is estimated for each of the 48 Fama-French industries (1997) with at least 20 firms in year *t*. This yields firm- and year-specific residuals that form the basis for the earnings value relevance metric: $VR_{i,t} = -[\ \sigma(\varepsilon_{i,t})]$ which is the standard deviation of firm *i's* residuals ($\varepsilon_{i,t}$) calculated over years *t-4* through *t* multiplied by negative one to obtain the following interpretation: the lower (higher) the information precision of a firm, the less (more) value relevant its earnings, the more (less) negative its standard deviations and *VR* metric, respectively. That is, *VR* is assumed to be positively associated with precision. All level variables are winsorized at the 1st and 99th percentile in each year.

Security Analyst Forecasts. Based on seminal work in Barron et al. (1998, BKLS), analyst forecasts can be used to make inferences about the underlying information environment of a stock. In particular, utilising observable characteristics of analyst forecasts (viz. forecast dispersion, squared error in the mean forecast and the number of forecasts), the BKLS model allows for an inference about the degree of information precision of the public and private information sets available to security analysts (see Eq. (3.13) and (3.14) for calculations). The theoretical validity of theses proxies is established by the fact that precision measures based on analyst forecasts substitute well for the general information precision of sophisticated investors (Barron et al., 2005).[70] Consistent with Botosan and Plumlee (2013), I use the sum of *public* and *private* as an analyst forecast-based indicator for *total* information precision.

$$public = \frac{(SE - \frac{D}{N})}{\left[\left(SE - \frac{D}{N}\right) + D\right]^2} \tag{3.13}$$

$$private = \frac{D}{\left[\left(SE - \frac{D}{N}\right) + D\right]^2} \tag{3.14}$$

where SE = squared error in the mean forecast $(\bar{F}_{it} - A_{it})^2$; D = forecast dispersion $\frac{1}{N-1}\sum_{i=1}^{N}\left(\bar{F}_{it} - F_{ijt}\right)^2$; N = number of forecasts; $\bar{F}_{it}$ = mean forecast for firm *i* in quarter *t*; A_{it} = actual earnings for firm *i* in quarter *t* and F_{ijt} = analyst *j*'s forecast of earnings for firm *i,* quarter *t. SE*, *D* and *N* are obtained from *Thomson Reuters I/B/E/S* summary files for the most recent one-quarter ahead earnings forecasts. For instance, if the forecast period is June 30th (q2) and there are forecast for q2 in Jan, Feb, March and April available, the forecasts given in April for period q2 are to be used. Further, only those observations are kept for which earnings are announced within 90 days after the quarter end. I also exclude all firm-quarter observations for which *D* is zero (as I interpret zero variance in forecasts as indication of stale forecasts) and a minimum of two unique analyst forecasts are required to be included in the sample.

The final measure is a time-series average of the quarterly values for *public* and *private* over the three quarters q-1 to q+1, where q equals the quarter in which the fiscal year

---

[70] See Barron et al. (2002) for an in-depth explanation of the BKLS model. Sheng and Thevenot (2012) propose a modification of the BKLS-based measures; they use a GARCH model to estimate the variance of the errors in the mean forecasts which they then substitute for the SE in the original BKLS-model. This method reduces measurement error and increase significance of statistical tests. However, the downside is that this approach requires extensively long firm-specific time-series data compared to the original approach. Thus, I choose to have a larger sample over the benefit of more significant results.

ends.[71] Following this method, for each firm-year, an average public and private precision measure is obtained. Given that *Precision* in the analytical models is the inverse of the variance of the private and public information signals, negative values of public and private are not meaningful and, thus, disregarded. The sum of *public* and *private* can then be thought of as an analyst forecast-based indicator for *total* information precision.[72] It is hypothesised that higher (lower) information precision of a firm is indicated by higher (lower) values of total analyst forecast precision.

Appendix 3.6: Empirical Measures of Information Asymmetry

Information asymmetry is measured by two microstructure indicators (i.e., probability of informed trading scores and bid/ask spreads) and one indicator for market competition (i.e., investor concentration). The measurement of each indicator is described below and all of them are conjectured to be positively associated with information asymmetry.

Market Microstructure. The annual average of simple daily percentage spreads is used as an indirect measure of information asymmetry, and probability of informed trading (PIN) scores are employed as a direct one.

Bid-Ask Spreads. Copeland and Galai (1983) and Glosten and Milgrom (1985) formally show that bid-ask spreads are valid measures for the exposure of market makers to the adverse selection problem, and as such capture well the degree of information asymmetry between informed and uninformed investors. While some authors use computationally demanding proxies for the adverse selection component of bid-ask spreads; I use annual averages of simple daily percentage spreads as a valid alternative (e.g. Coller and Yohn

---

71 As an alternative, Botosan and Plumlee (2013) takes the average over the four quarters q-3 to q. Following this latter approach leaves my SEM results unaffected, but reduces sample size for AFP_Total by about 1,000 firm-years.

72 Note that there is a separate literature that exclusively uses *analyst forecast dispersion* as a proxy for information precision (e.g., Diether et al. (2002)), however, dispersion represents only one element of uncertainty (i.e., uncertainty arising from analysts' private information), thus, BKLS-based measures are to be preferred as they account for both public (common) and private (idiosyncratic) uncertainty (Sheng and Thevenot, 2012, p.21)

(1997), Stoll (1978)).[73] Consistent with Stoll (1978), for each firm *i* the daily percentage spread is estimated as shown in Eq. (3.15). I then calculate the final SPREAD statistic for firm *i* as the average daily percentage spread over 252 trading days which are centred around the fiscal year end date (i.e., the average over daily spreads from t-126 to t+125 is taken). In case a firm's fiscal year end falls on a day of no trading, I use the latest available SPREAD statistic which immediately precedes the fiscal year end date.

$$spread = \frac{|P^B - P^A|}{\frac{1}{2}(P^B + P^A)} \tag{3.15}$$

where $|P^B - P^A|$ = absolute difference between closing bid (P^B) and closing ask (P^B) prices and $\frac{1}{2}(P^B + P^A)$ = mid-point of the bid/ask closing prices.

Probability of Informed Trading Scores. Developed by Easley et al. (1996, 1997), PIN scores are firm-specific proxies for information asymmetry in that they measure the probability that the next trade order is from a privately informed investor, with larger PIN scores signifying larger information asymmetry. The underlying notion of the PIN model—shown in Eq. (3.16)—is that while it is impossible to directly observe which trades are based on private information, one can use imbalances between buy and sell orders to infer the probability of information-based trading for a given stock.[74]

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s} \tag{3.16}$$

where α = probability of a private information event; μ = the daily rate of informed trade arrival; $\varepsilon_b(\varepsilon_s)$ = daily rate of uninformed buy (sell) trade arrival.

However, PIN scores based on the original model tend to be a rather sensitive proxy for information asymmetry, given its difficulties to deal with large firms and high trading

---

73 Different methodologies can be followed in estimating the adverse selection component of bid-ask spreads; e.g. Bhattacharya et al. (2012, 2013) follow Huang and Stoll (1996); Armstrong et al. (2011) follow Madhavan et al. (1997); and Akins et al. (2012) follow Glosten and Harris (1988).

74 Brown et al. (2004, pp. 348-349) give a thorough description of the PIN model.

volumes. Easley et al. (2010, p. 300) state that PIN scores "provide a better description of the information environment among smaller stocks" given the conjecture of the model that information events occur only once per day, which seems to be a less reasonable assumption for larger firms. Further, the model is plagued by computational under-/overflow when dealing with extremely large trading volumes which nowadays are the norm rather than the exception. What is more, the model assumes that buy and sell orders from uninformed investors are uncorrelated, while in practice a positive association is often the case (e.g., release of macroeconomic information or earnings announcements increases trading of both uninformed buyers and sellers). Brown and Hillegeist (2007) address these shortcomings and provide a robuster PIN model that extends seminal work in Venter and De Jongh (2006).

Stephen Brown provides quarterly PIN scores for all CRSP listed firms from 1993 to 2010 on his website.[75] As recommended by Brown, I drop all firm-quarters for which the number of trading days is below thirty and calculate the final PIN measure as a time-series average of the quarterly PIN values over the three quarters q-1 to q+1, where q equals the quarter in which the fiscal year ends.

Market Competition. Consistent with Akins et al. (2012), I measure market competition by means of *investor concentration*, but utilise information on mutual fund holdings instead of institutional investor holdings due to data restrictions. *Thomson Reuters' Mutual Funds Holding (s12)* database provides security holding information for all registered mutual funds that report their holdings with the SEC. The calculation of the concentration measure (*INV_Conc*) is shown in Eq. (3.17). For each firm-year, the final *INV_Conc* measure is an average over the three quarters q-1 to q+1, where q = quarter in which the fiscal year ends and it is conjectured that higher values of concentration denote *less* competition in the trades of firm *i*'s stocks; hence, *INV_Conc* is assumed to be positively associated with information asymmetry.[76]

75 http://scholar.rhsmith.umd.edu/sbrown/pin-data

76 Akins et al. multiply the index by negative one, so that higher values indicate *more* competition (p. 41); however, in order to maintain consistent interpretation between all of my indicators and asymmetry (viz. a conjectured positive association), I refrain from this modification.

$$INV_{Conc} = \sum_{j=1}^{N} \left(\frac{Fund_{i,j}}{Fund_i} \right)^2 \quad (3.17)$$

where $Fund_{i,j}$ = number of shares held by mutual fund *j* in firm *i*, $Fund_i$ = total shares held by all mutual funds in firm *i*, and *N* = is the total number of mutual funds invested in in firm *i*.

Appendix 3.7: Empirical Measures of Cost of Equity

The proceeding paragraphs describe the estimation of the risk factor based (RFB) and valuation model based (VMB) expected return proxies used in this study.

Risk factor-based CoE proxies. A common method to estimate expected return is to use asset pricing models. For instance, the Arbitrage Pricing Theorem (ATP) of Ross (1976) states that the expected rate of return of a firm is equal to the risk free rate ($r_{f,t}$) plus the sum of the risk premia ($er_{k,t} - r_{f,t}$) of *K* risk factors multiplied by the firm's respective sensitivity to these factors (β_k)—that is, the CoE of a firm is a linear function of its exposure to a number of different risk factors—see Eq. (3.18). Given this insight, I estimate three widely used risk factor based CoE proxies as described below.

$$er_t = r_{f,t} + \sum_{k=1}^{K} \beta_k \left(er_{k,t} - r_{f,t} \right) \quad (3.18)$$

Carhart's Four-Factor Model (rFF4). One of the most widely used CoE measures is obtained from Carhart's (1997) four-factor model (r_{FF4}) that extends Fama and French's (1993) three-factor model by a momentum factor which captures the return differences between positive and negative momentum stocks (Jegadeesh and Titman, 1993). In line with Barth et al. (2013), I calculate each firm's cost of equity (r_{FF4}) for year *t+1* as of year *t* as shown in equation (3.19). This involves the following steps.

$$\begin{aligned} r_{FF4} = \bar{r}_{f,t} &+ \hat{\beta}_{RMRF,i,t} \left(\overline{R_M - R_f} \right)_t + \hat{\beta}_{SMB,i,t} \overline{(SMB)}_t + \hat{\beta}_{HML,i,t} \times \overline{(HML)}_t \\ &+ \hat{\beta}_{UMD,i,t} \times \overline{(UMD)}_t \end{aligned} \quad (3.19)$$

First, I collect monthly time-series returns on the Fama-French and momentum factors from Kenneth French homepage and WRDS database, respectively, and calculate expected annual factor returns (viz. $\left(\overline{R_M - R_f}\right)_t$ / $\overline{SMB}_t$ / $\overline{HML}_t$ / $\overline{UMD}_t$) by taking each factor's average monthly return over rolling 60 month windows.[77] Monthly average returns are then annualised by compounding over 12 months. Similarly, the expected annual risk free rate ($\bar{r}_{f,t}$) is obtained, however, 12 month rolling windows instead of 60 month windows are used to avoid outdated estimates (Barth et al. (2013), Landsman et al. (2011)).

Second, I estimate Eq. (3.20) by means of rolling time-series regressions over 60 month windows to attain firm-specific end of month factor loadings (i.e., $\hat{\beta}_{RMRF,i,t}$; $\hat{\beta}_{SMB,i,t}$; $\hat{\beta}_{HML,i,t}$; $\hat{\beta}_{UMD,i,t}$). Finally, I calculate each firm's CoE (r_{FF4}) six months after its fiscal year end by evaluating Eq. (3.19) to ensure that end of year information has been impound in stock prices.

$$\begin{aligned} RET_{i,m} - r_{f,m} = {} & \alpha_i + \beta_{RMRF,i,m}\left(R_M - R_f\right)_m + \beta_{SMB,i,m}(SMB)_m \\ & + \beta_{HML,i,m} \times (HML)_m + \beta_{UMD,i,m} \times (UMD)_m \end{aligned} \tag{3.20}$$

where $RET_{i,m}$ = firm's i total monthly stock return; $r_{f,m}$ = risk-free interest rate (one-month treasury bill rate); $\left(R_M - R_f\right)$ = monthly excess return on the market; $(SMB)_m$, $(HML)_m$ and $(UMD)_m$ = monthly returns on the Fama-French and momentum factors.

Fama-French Three-Factor (rFF3) and CAPM model (rCAPM). Following the same methodology as for rFF4, I estimate rFF3 and rCAPM by evaluating Eq. (3.21) and (3.22).

$$r_{FF3} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t + \hat{\beta}_{SMB,i,t}\overline{(SMB)}_t + \hat{\beta}_{HML,i,t} \times \overline{(HML)}_t \tag{3.21}$$

$$r_{CAPM} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t \tag{3.22}$$

[77] An alternative approach would be to follow Kothari et al. (2009) and simply average the monthly factor returns over the entire time-series (in their study from 1963-2000). However, this tends to induce forward looking bias in the calculations and assumes time-invariant factor loadings.

Risk factor-based CoE proxies (FVIX augmented). Consistent with Ang et al. (2006), I re-estimate the three RFB models with an additional risk-factor for expected market volatility (i.e., FVIX). FVIX reflects the monthly excess return on a factor-mimicking portfolio that tracks daily changes in the VIX index (Barinov, 2013, p. 1880). The notion underlying the FVIX factor is that companies with more (less) negative return sensitivity to VIX index changes have higher (lower) CoE. Following the same methodology as for rFF4, I estimate rFVIX4, rFVIX3 and rFVIX by evaluating Eq. (3.23) to (3.25).

$$r_{FVIX4} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t + \hat{\beta}_{SMB,i,t}\overline{(SMB)}_t + \hat{\beta}_{HML,i,t} \times \overline{(HML)}_t + \hat{\beta}_{UMD,i,t} \times \overline{UMD)}_t + \hat{\beta}_{FVIX,i,t} \times \overline{(FVIX)}_t \tag{3.23}$$

$$r_{FVIX3} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t + \hat{\beta}_{SMB,i,t}\overline{(SMB)}_t + \hat{\beta}_{HML,i,t} \times \overline{(HML)}_t + \hat{\beta}_{FVIX,i,t} \times \overline{(FVIX)}_t \tag{3.24}$$

$$r_{FVIX} = \bar{r}_{f,t} + \hat{\beta}_{RMRF,i,t}\left(\overline{R_M - R_f}\right)_t + \hat{\beta}_{FVIX,i,t} \times \overline{(FVIX)}_t \tag{3.25}$$

Valuation model-based CoE Proxies. I estimate the four different ICC measures suggested by Easton (2004) which are all based on the abnormal earnings growth model (see Eq. (3.26) - (3.28)). The models differ with respect to assumption made about dividend payments, short-term and perpetual abnormal earnings growth. The simplest ICC estimate (rPE) assumes no dividend payout nor any short- or perpetual growth in abnormal earnings. In contrast, rAEGM makes use of all available assumptions, with rPEG and rMPEG lying between those two “extreme” cases.[78]

Price-Earnings-Ratio (rPE)

$$r_{PE} = \left(\frac{P}{eps_1}\right)^{-1} \tag{3.26}$$

[78] See Echterling et al. (2015) for a recent review of the literature on the computation and assessment of different implied cost of capital measures.

where *eps1* = one year ahead I/B/E/S earnings consensus forecasts and *P* = current price from I/B/E/S pricing files. All observations for which *eps1* is negative are excluded.

Price-Earnings-Growth (rPEG)

$$r_{PEG} = \sqrt{\frac{eps_2 - eps_1}{P}} \tag{3.27}$$

where *eps1* (*eps2*) = one (two) year ahead I/B/E/S earnings consensus forecasts and *P* = current price from I/B/E/S pricing files. All observations for which *eps1* is larger than *eps2* are excluded.

Abnormal Earnings Growth (rAEGM)

$$r_{AEGM} = A + \sqrt{A^2 + \frac{eps_1}{P}\left(\frac{eps_2 - eps_1}{eps_1} - G_{AEG}\right)} \tag{3.28}$$

where $A = \frac{1}{2}\left(G_{AEG} + \frac{dps_1}{P}\right)$; G_{AEG} = perpetual growth rate in abnormal earnings set to current expected annual risk-free rate minus three percent. The expected annual risk-free rate is calculated by first taking the average of the one-month Treasury bill rate over rolling past 12-month windows and then annualise it by compounding over 12 months. Risk-free rates are from Kenneth French homepage. *dps1* = analysts' mean dividends per share estimates for t+1 as reported on IBES summary files; if missing, *dps1* is forecasted as *eps1* x *current dividend payout ratio*. *Dividend payout ratio* is calculated as $\frac{Dividends\ (Common)}{Income\ Before\ Extraordinary\ Items}$ for firms with positive earnings and $\frac{Dividends\ (Common)}{0.06\ x\ Total\ Assets}$ for firms with current negative earnings. *Eps1*, *eps2* and *P* as describe above. Current accounting data is from Compustat.

Appendix 3.8: Median ROA (%) by Constructs & Indicators

	Quantity			Precision			Asymmetry		
Fiscal Year	**Listing**	**Age**	**RMC**	**AQ**	**VR**	**AFP_Total**	**PIN**	**SPREAD**	**INV_Conc**
1993	4.19	5.19	4.92	4.23	3.71	4.08	3.71	3.11	3.52
1994	4.75	5.71	5.21	4.92	4.19	4.75	4.34	3.74	4.08
1995	5.11	5.67	5.59	5.20	4.36	5.14	4.49	3.85	4.07
1996	5.09	6.04	5.77	5.27	4.39	4.76	4.53	3.82	4.07
1997	5.09	5.98	5.72	5.10	4.33	4.47	4.21	3.91	4.24
1998	4.23	5.54	4.97	4.42	3.44	3.75	3.44	3.30	3.57
1999	4.21	4.94	4.92	4.60	3.44	4.04	3.49	3.37	3.58
2000	3.93	4.84	4.85	4.53	3.39	4.28	3.39	3.06	3.52
2001	2.40	3.57	3.48	3.04	2.00	2.59	2.07	1.79	2.10
2002	2.63	3.53	3.61	3.15	2.30	2.70	2.31	2.12	2.38
2003	2.97	3.95	4.02	3.57	2.76	3.28	2.68	2.52	2.74
2004	3.71	4.71	4.81	4.67	3.47	4.36	3.50	3.32	3.49
2005	4.10	5.09	5.24	4.81	3.76	4.51	3.85	3.66	4.06
2006	4.20	5.33	5.47	4.73	3.66	4.62	3.87	3.68	4.08
2007	4.01	5.25	5.36	4.86	3.46	4.62	4.00	3.59	4.02
2008	2.31	3.74	3.85	3.40	1.85	2.93	2.41	1.87	2.41
2009	1.48	2.52	2.84	2.58	1.50	2.47	1.73	1.46	1.73
2010	3.00	4.16	4.50	4.40	2.94	3.87	4.27	2.93	3.31
Average	**3.75**	**4.76**	**4.73**	**4.30**	**3.28**	**3.96**	**3.46**	**3.06**	**3.39**

Table continued on next page.

Appendix 3.8: Median ROA (%) by Constructs & Indicators (cont.)

Sample ROA	Valuation Model-Based CoE			Risk Factor-Based CoE			Future Realised Returns			
Fiscal Year	**rPE**	**rPEG**	**rAEGM**	**rCAPM**	**rFF3**	**rFF4**	**ret12**	**ret24**	**ret36**	**Market ROA**
1993	3.90	3.94	3.96	3.06	3.06	3.06	3.17	3.32	3.49	1.96
1994	4.38	4.71	4.71	3.73	3.73	3.73	3.82	3.90	4.02	2.35
1995	4.50	4.69	4.71	3.78	3.78	3.78	3.90	4.02	4.12	2.32
1996	4.49	4.40	4.40	3.85	3.85	3.85	3.93	4.03	4.12	2.29
1997	4.57	4.53	4.52	3.89	3.89	3.89	4.07	4.13	4.24	2.08
1998	4.03	3.82	3.81	3.29	3.29	3.29	3.37	3.41	3.56	1.60
1999	4.13	4.14	4.16	3.35	3.35	3.35	3.39	3.46	3.51	1.44
2000	4.18	4.22	4.22	3.12	3.12	3.12	3.14	3.20	3.30	1.07
2001	3.14	2.67	2.67	1.78	1.78	1.78	1.85	1.98	2.06	0.74
2002	3.20	2.86	2.88	2.09	2.09	2.09	2.19	2.29	2.31	1.14
2003	3.55	3.33	3.33	2.48	2.48	2.48	2.61	2.68	2.74	1.50
2004	4.31	3.83	3.83	3.28	3.28	3.28	3.34	3.45	3.49	2.29
2005	4.66	4.33	4.37	3.66	3.66	3.66	3.72	3.80	3.82	2.36
2006	4.71	4.41	4.38	3.71	3.71	3.71	3.76	3.86	3.91	2.60
2007	4.84	4.44	4.46	3.66	3.66	3.66	3.70	3.79	3.88	2.17
2008	3.82	2.81	2.82	1.90	1.90	1.90	1.95	2.04	2.04	0.77
2009	2.88	2.10	2.14	1.48	1.48	1.48	1.53	1.56	1.65	0.97
2010	4.22	3.56	3.60	2.95	2.95	2.95	2.98	3.04	3.07	2.46
Average	**4.08**	**3.82**	**3.83**	**3.06**	**3.06**	**3.06**	**3.13**	**3.22**	**3.29**	**1.78**

Market ROA is the median ROA (Income Before Extraordinary Items / Total Assets) of all CMM listed firms from 1993 to 2010.

Appendix 3.9: Market Capitalisation (% of Total Market Cap) by Constructs & Indicators

	Quantity			Precision			Asymmetry			
Fiscal Year	**Listing**	**Age**	**RMC**	**AQ**	**VR**	**AFP_Total**	**PIN**	**SPREAD**	**INV_Conc**	**Total Mkt. Cap**
1993	24.72	23.50	30.87	24.76	58.61	22.82	63.71	66.57	43.08	$6,295bn
1994	24.61	23.29	30.47	43.53	57.36	23.12	63.48	66.56	43.15	$6,512bn
1995	25.08	24.56	31.49	44.29	58.06	22.57	64.68	65.78	43.45	$8,534bn
1996	25.27	24.42	31.29	42.51	55.96	21.50	64.17	64.88	45.44	$10,520bn
1997	28.01	26.52	33.42	42.88	57.59	24.03	66.51	66.42	47.28	$13,741bn
1998	29.33	27.68	35.31	45.24	59.19	25.41	68.22	67.87	47.33	$16,511bn
1999	26.85	24.65	33.04	44.64	55.50	26.76	62.93	62.24	47.98	$22,086bn
2000	33.14	29.82	37.76	47.62	58.93	29.78	66.85	58.37	52.41	$21,486bn
2001	33.17	30.45	38.65	48.22	61.06	36.69	68.18	68.43	54.69	$18,441bn
2002	33.90	31.27	38.73	48.81	62.35	39.12	69.56	69.89	58.39	$14,742bn
2003	33.52	30.27	37.21	47.90	61.52	40.99	68.12	68.12	58.45	$19,679bn
2004	34.84	31.24	37.33	46.30	61.63	42.21	65.98	68.74	49.09	$22,201bn
2005	33.62	30.60	36.76	46.55	60.42	41.07	67.16	67.87	59.23	$23,658bn
2006	33.54	31.19	36.34	46.95	61.38	41.66	68.34	69.62	61.89	$27,300bn
2007	34.87	33.55	38.61	53.02	65.63	43.66	44.39	73.45	65.76	$28,160bn
2008	37.57	37.42	42.77	56.30	68.56	49.22	45.92	76.94	71.05	$17,368bn
2009	37.08	36.55	41.98	56.46	69.51	50.99	72.88	77.39	71.18	$21,759bn
2010	39.40	38.64	43.39	56.67	72.93	50.41	7.61	77.30	72.31	$24,656bn
Average	**31.58**	**29.76**	**36.41**	**46.81**	**61.45**	**35.11**	**61.04**	**68.69**	**55.12**	**323,647.00**

Table continued on next page.

Appendix 3.9: Market Capitalisation (% of Total Market Cap) by Constructs & Indicators (cont.)

	Valuation Model-Based CoE			Risk Factor-Based CoE			Future Realised Returns			
Fiscal Year	**rPE**	**rPEG**	**rAEGM**	**rCAPM**	**rFF3**	**rFF4**	**ret12**	**ret24**	**ret36**	**Total Mkt. Cap**
1993	43.25	41.82	41.80	68.66	68.66	68.66	67.65	66.03	65.10	$6,295bn
1994	44.65	41.44	41.41	68.44	68.44	68.44	66.83	66.08	64.28	$6,512bn
1995	44.42	40.63	40.45	67.82	67.82	67.82	67.13	65.27	61.86	$8,534bn
1996	46.18	43.76	43.43	67.01	67.01	67.01	65.20	61.79	59.19	$10,520bn
1997	48.07	45.02	44.78	68.63	68.63	68.63	65.32	62.76	58.70	$13,741bn
1998	52.81	51.00	50.94	70.34	70.34	70.34	67.62	63.57	62.40	$16,511bn
1999	48.38	45.63	45.59	64.73	64.73	64.73	61.71	60.75	60.15	$22,086bn
2000	50.89	46.57	46.07	69.53	69.53	69.53	68.67	67.86	66.94	$21,486bn
2001	54.22	53.27	53.22	71.08	71.08	71.08	70.49	69.45	68.60	$18,441bn
2002	57.92	50.98	50.98	72.58	72.58	72.58	71.48	70.49	68.52	$14,742bn
2003	57.09	51.68	51.68	70.78	70.78	70.78	69.77	67.61	65.99	$19,679bn
2004	59.22	51.31	51.29	71.63	71.63	71.63	69.39	67.64	66.48	$22,201bn
2005	59.80	52.14	52.12	71.03	71.03	71.03	69.07	67.88	66.55	$23,658bn
2006	61.44	56.51	53.78	72.98	72.98	72.98	72.07	70.73	69.95	$27,300bn
2007	65.48	63.09	62.46	76.72	76.72	76.72	75.76	75.03	74.10	$28,160bn
2008	69.25	64.44	64.43	80.24	80.24	80.24	79.48	78.51	78.10	$17,368bn
2009	71.14	68.62	68.51	80.62	80.62	80.62	79.77	79.25	78.80	$21,759bn
2010	72.35	67.59	67.55	80.71	80.71	80.71	80.29	79.80	79.09	$24,656bn
Average	**55.92**	**51.97**	**51.69**	**71.86**	**71.86**	**71.86**	**70.43**	**68.92**	**67.49**	**323,647.00**

Total Market Capitalisation in billion USD which represents market value (common shares outstanding x fiscal year end closing stock price) for all CMM listed firms from 1993 to 2010.

4 Implied Cost of Capital and Cross-Sectional Earnings Forecasting Models: Evidence from Newly Listed Firms

Abstract

Elaborating on seminal work by Hou et al. (J Account Econ, 53:504-526, 2012, HVZ) and Li and Mohanram (Rev Account Stud, 19:1152-1185, 2014, LM) on how cross-sectional earnings forecasting models can address prevailing limitations in implied cost of capital (ICC) research, this study evaluates mechanical earnings predictions for newly listed firms in terms of forecast bias, forecast accuracy and earnings response coefficients (ERC). Using three cross-sectional earnings forecasting models suggested by HVZ (2012) and LM (2014), I provide for a sample of 1,657 IPOs comparative evidence on the quality of model-based earnings forecasts and the validity of ICC estimates derived therefrom. Results demonstrate that combining the earnings model of HVZ (2012) with the earnings persistence (EP) model of LM (2014) into one forecasting solution (HVZ/EP) generates less forecast bias, higher ERCs and more valid ICC estimates vis-à-vis the HVZ, EP and RI (residual income) models stand-alone. This suggests that for smaller and younger firms more complex forecasting solutions might be required to ensure reliability of model-based earnings predictions and ICC calculations. The average IPO in my sample has an implied cost of capital of 10.9 percent, which is consistent with model- and analyst-based ICC estimates in extant research.

4.1 Introduction

The expected rate of return of a firm's equity (also referred to as cost of equity, discount rate or required return) is one of the most crucial numbers in corporate finance: investors require reliable estimates for equity valuation, firm managers need it for capital budgeting and academics use it as a dependent variable in a variety of settings. While its appropriate determination is an ongoing debate, its general influence on corporations is clear: lower costs of equity (CoE), lead to higher valuations (i.e., higher stock prices) which is tantamount to increased shareholder wealth.

Traditional risk-factor based (RFB) cost of equity measures (e.g., CAPM, FF3-factor estimates) rely on past realised returns to gauge firms' expected returns. However, the unexpected news component in realised returns (i.e., noise) tend to corrupt the reliability of factor loading and factor premia estimates within asset pricing models, which may lead to "woefully imprecise estimates of the cost of equity" (Fama and French, 1997, p. 154). Given this shortcoming, an aspiring literature—in particular in accounting research—applies the concept of implied cost of capital (ICC) to derive alternative proxies for expected rate of returns, which I refer to as valuation model-based (VMB) proxies in this study. The applicability of the ICC methodology crucial depends on the availability of reliable predictions of future payoffs to shareholders and extant reserach primarily uses analysts' consensus earnings forecasts to proxy for them. However, with the reliance on these forecasts two problems arise.

First, analyst-based earnings forecasts tend to be overly optimistic (e.g., Francis and Philbrick (1993), Dugar and Nathan (1995), McNichols and O'Brien (1997)) which may result in biased and invalid ICC estimates. For instance, Easton and Monahan (2005, p. 501) investigate seven different analyst-based ICC estimates and conclude that "for the entire cross-section of firms, these proxies are unreliable", but also confirm that some proxies become reliable when "analysts' forecast accuracy is high". Second, young, small and financially distressed firms are rarely covered by security analysts (e.g., Diether et al. (2002), Hong et al. (2000), La Porta (1996)) which limits the ICC methodology towards large and well-established firms. This coverage bias makes analyst-based ICC measures less of an alternative to RFB proxies in that the latter only require firms to have a sufficiently long return history to be estimated.

This paper complements seminal work by Hou, van Dijk and Zhang (2012, hereafter: HVZ) and Li and Mohanram (2014, hereafter: LM) on how cross-sectional earnings forecasting models can address analyst-associated deficiencies in ICC research.[79] In particular, I am testing the applicability of the HVZ, EP and RI model for the smallest, youngest and least followed firms in equity markets: initial public offerings (IPOs). The main objective of this paper is to investigate as to what extent the respective models can be used to (1) predict earnings and (2) derive implied cost of capital of newly listed firms. More specifically, I examine if earnings forecasts based on pre-IPO financial information can be used to derive valid ICC estimates for those firms.

For a sample of 1,657 IPOs from 1972-2013, I find that combining the HVZ and EP model into one forecasting solution (HVZ/EP) outperforms earnings predictions from the HVZ, EP and RI model stand-alone, as indicated by its lower forecast biases and higher earnings response coefficients (ERCs). The average (median) forecast bias form the HVZ/EP model is -1.61 (-0.32) percent of market value of equity and average (median) forecast accuracy equals 6.11 (3.73) percent. These levels are consistent with large sample study figures in HVZ (2012) and LM (2014), and endorse the general applicability of earnings forecasting models for smaller, younger and less followed firms. From the earnings forecasts, I calculate for each model and IPO a composite ICC (rCOMP) at seven different points in time after initial listing (using closing price at the end of one, three, six, nine, 12, 18 and 24 month of trading).[80] Expected returns for the composite ICC are highly consistent across all four models and over time. The average (median) IPO in my sample has an expected rate of return of 10.9 (8.6) percent, which is consistent with analyst-based ICC estimates in Liu et al. (2014) for a sample of approximately 800 IPOs.

[79] HVZ (2012) predict earnings from contemporary accounting data (viz. total assets, dividends, earnings and accruals) and LM (2014) propose two parsimonious alternatives to the HVZ model: the earnings persistence (EP) and residual income (RI) model.

[80] The composite ICC (rCOMP) is the average of rPE, rPEG, rMPEG and rAEGM (Easton, 2004).

To assess the construct validity of my model-based ICC estimates, I regress one-year ahead buy-and-hold returns on each models composite ICC. Fama-MacBeth (FMB) coefficients for HVZ, EP, RI and HVZ/EP are all negative and insignificant when ICCs are calculated from first month closing prices (i.e., MV1), somewhat mixed (7 negative; 5 positive) for MV3 to MV9, but all positive and significant for forecasting horizons beyond (minimum t-stat. 1.80). Inefficient equity valuations, in the form of upward biased share prices in the immediate IPO aftermarket, might explain the initially insignificant results. What is more, HVZ/EP-based estimates display the highest construct validity across all models in that FMB coefficients are consistently close to one from MV12 onwards (MV12: 0.94; MV18: 1.04; MV24: 0.96), whereas coefficients for HVZ, EP and RI are above one for MV12 (1.31; 1.22; 1.12) and below one for MV24 (0.87; 0.74; 0.75), indicating that these estimates are on average too low for MV12 and too high for MV24.

Overall, findings show that the quality of forecasted IPO earnings and validity of ICC estimates derived therefrom is highest for the HVZ/EP model, followed by HVZ in second, and EP & RI in joint-third place. This implies that when predicting earnings for smaller and younger firms model performance decreases with parsimony and, thus, more complex forecasting solutions seem required (such as HVZ/EP) for those types of firms.

The rest of this paper proceeds as follows. Section 4.2 reviews the related literature and the earnings forecasting models employed in this study. Section 4.3 outlines the methodology. Section 4.4 evaluates the performance of the model-based earnings forecasts and ICC estimates. Concluding remarks are provided in Section 4.5.

4.2 Related Literature

4.2.1 The Association between Realised and Expected Returns

Based on Campbell and Shiller (1988a, 1988b), firms' realised returns at date *t+1* can be decomposed into an expected and unexpected return component. Formally:

$$r_{t+1} = er_t + \delta_{t+1} \tag{4.1}$$

where r_{t+1} is the firm's realised return at date $t+1$, er_t is the expected rate of return (CoE) at the beginning of $t+1$ conditional on all available information at date t and δ_{t+1} captures the unexpected (abnormal) return component from date t to $t+1$. The unexpected return component is due to new information and can be further decomposed into unexpected returns due to (1) cash flow news and (2) discount news; that is *unexpected* stock price movements (i.e., abnormal returns) occur either because there are surprising idiosyncratic cash flow news and/or because the underlying systematic risk of an asset changes. In equation form:

$$r_{t+1} = er_t + (cn_{t+1} - rn_{t+1}) \tag{4.2}$$

where cn_{t+1} is return due to cash flow news and rn_{t+1} is return due to discount rate news from date t to $t+1$, respectively. The negative sign on discount rate news ($-rn_{t+1}$) captures the fact that, ceteris paribus, an increase in future discount rates lead to a decrease in stock price, resulting in realised returns being lower than expected returns.

It follows from equation (4.1) and (4.2) that while in hindsight realised returns can be explained by both its expected and unexpected component, the *ex ante* predictability of realised returns is only due to the expected return component; that is, er_t is the only "statistical object that predicts returns" (Lee et al., 2010, p. 7). Although the true expected return of a firm is not observable and therefore subject to measurement error, the insight is clear: the more precise the measure for expected returns, the more precise the return forecasts. In general, the literature takes two main approaches when measuring firms expected returns; the first approach relies on past realised returns—which I refer to as risk factor-based (RFB)—and the second one conforms to the concept of implied cost of capital—denoted as valuation model-based (VMB) in this paper. Table 4.1 provides an overview of commonly applied CoE measures in extant work.

Table 4.1: Risk Factor-Based vs. Valuation Model-Based CoE proxies

Risk Factor-Based (RFB)	Valuation Model-Based (VMB)
▪ **Asset Pricing Models** - CAPM: r_{CAPM} - (Lintner, 1965, Mossin, 1966, Sharpe, 1964) - Fama-French three factor model: r_{FF3} - (Fama and French, 1993) - Carhart's four-factor model: r_{FF4} - (Carhart, 1997) - Fama-French five factor model: r_{FF5} - (Fama and French, 2015)	▪ **Dividend Discount Model** - Finite horizon: r_{GOR} - (Gordon and Gordon, 1997) - Target price: r_{BP} - (Botosan and Plumlee, 2002) ▪ **Residual Income Model** - Economy-wide: r_{CT} - (Claus and Thomas, 2001) - Industry method: r_{GLS} - (Gebhardt et al., 2001) ▪ **Abnormal Earnings Growth Model** - Price-earnings ratio: r_{PE} - (Easton, 2004) - Price-earnings-growth ratio: r_{PEG} - (Easton, 2004) - Modified price-earnings growth: r_{MPEG} - (Easton, 2004) - Change in abnormal earnings growth: r_{AEGM} - (Easton, 2004) - Economy-wide growth: r_{OJN} - (Ohlson and Juettner-Nauroth, 2005) - Modified economy-wide growth: r_{GM} - (Gode and Mohanram, 2003)

The table reports commonly used risk factor-based and valuation model-based CoE measures. Reported in parentheses are the original sources suggesting the respective proxy.

4.2.1.1 Risk Factor-Based Proxies

A time-honoured approach in estimating firms' CoE is to use past realised returns. One method to gauge the expected rate of return of a firm is to assume that the unexpected return component (δ_{t+1}) is mean zero over time and firms, which makes average past realised returns a natural proxy for future expected returns. This approach has been taken by some researchers (e.g., Doukas et al. (2006), Konchitchki et al. (2016)), however, it seems problematic to the extent that in particular firm-specific cash-flow news (cn_{t+1}) can have a significant impact on firm-level returns (e.g., Chen et al. (2013), Ogneva

(2012), Vuolteenaho (2002)).[81] A more common method is to directly estimate expected returns (er_t) by means of asset pricing models, where it assumed that firms' CoE equals the risk-free rate ($r_{f,t}$) plus the sum of several different risk-premia—formally:

$$er_t = r_{f,t} + \sum_{k=1}^{K} \beta_k \left(er_{k,t} - r_{f,t}\right) \tag{4.3}$$

However, irrespective of which asset pricing model is used (e.g., Carhart (1997), Fama and French (1993), Fama and French (2015), Jegadeesh and Titman (1993), Lintner (1965), Mossin (1966), Pástor and Stambaugh (2003), Sharpe (1964)), all of them are plagued by the fact that their estimates are based on noisy past realised returns; that is, the unexpected news component in realised returns tend to corrupt the reliability of factor loading and factor premia estimates in RFB models, which can result in "woefully imprecise estimates of the cost of equity" (Fama and French, 1997, p. 154).

4.2.1.2 Valuation Model-Based Proxies

Given the limitations of RFB measures, an aspiring literature—in particular in accounting research—follows the approach of implied cost of capital (ICC) in estimating firms' CoE (e.g., Bhattacharya et al. (2012), Botosan et al. (2004), Francis et al. (2004)). The intuition behind the ICC framework is straightforward: use a specific valuation model, accept the current stock price as at least semi-strong efficient in the classical efficient market hypotheses sense (Fama, 1965, 1970) and *back-out* the internal rate of return which equates current stock price of the firm with its expected future payoffs to shareholders. The internal rate of return is then considered as market participants' *ex ante* assessment of the firm's CoE. More formally, the discount rate (r) that equates P_{t-1} and $\sum_{t=1}^{\infty} E_{t-1}[dps_t]$ in the well-known dividend discount model—Eq. (4.4) —is regarded as a firm's expected rate of return.

[81] Elton (1999, p. 1199) also questions the reliability of average realised returns as proxies for expected returns: "The use of average realized returns as a proxy for expected returns relies on a belief that information surprises tend to cancel out over the period of a study and realized returns are therefore an unbiased estimate of expected returns. However, I believe that there is ample evidence that this belief is misplaced."

$$P_{t-1} = \sum_{t=1}^{\infty} (1+r)^{-t} E_{t-1}[dps_t] \tag{4.4}$$

where P_{t-1} is the current stock price, $E_{t-1}[dps_t]$ is expected future dividends conditional on all available information at time $t=0$ and r is the discount rate.

The residual income and abnormal earnings growth model are important alternatives to the dividend discount model and the ICC methodology can be easily extended towards these models (Claus and Thomas (2001), Gebhardt et al. (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005)).[82]

4.2.1.3 Performance Evaluation

The literature suggests three methods to assess the construct validity of the different CoE measures: first, comparison of the association between the proxy and future realised returns;[83] second, comparison of the association between the proxy and common risk factors (e.g., market beta, leverage, B/M, volatility, size);[84] and third, comparison of the predictive power of the proxy for future realised returns.

Studying eleven ICC estimates and one RFB proxy in regard to their association with common risk factors and explanatory power of future realised returns, Botosan et al. (2011) find that the ICC measures based on price-earnings-growth-ratios within an abnormal earnings growth model (rPEG) and target prices within a standard dividend discount model (rBP) demonstrate greatest construct validity among all proxies tested. In a similar vein, Lee et al. (2010) examine the predictive power of seven ICC and two RFB proxies and conclude that “all of the ICC estimates tested perform much better than the beta-based measures widely touted in finance textbooks” (ibid., p. 26).[85] Taken findings

82 See also Echterling et al. (2015) who provide a recent review of alternate methods of computation.

83 Two prominent examples following this methodology are Guay et al. (2011) who regress realised returns on different ICC proxies; and Easton and Monahan (2005) who extend Guya et al. in that they control for cash flow and discount rate news in their regression.

84 Botosan and Plumlee (2005) follow this approach to assess the validity of different ICC proxies. Easton and Monahan (2016) argue that this approach is illogical, given that the ICC approach is based on the assumption that these very risk factors are either unknown or cannot be measured reliably.

85 Lee et al. (2015) report similar results in both cross-sectional and time-series analyses.

together, this suggest that VMB proxies are more valid estimates of firms' expected returns than RFB ones.

4.2.2 Earnings Forecasting Models

Despite the compelling evidence of VMB estimates demonstrating greater construct validity than traditional RFB measures, the ICC methodology is not impeccable either. In particular, ICC estimates only yield an unbiased CoE estimate if market prices are efficient and forecasted future payoffs are congruent with overall market expectations. Assuming markets are at least semi-strong efficient—an assumption which seems to be violated in the immediate IPO aftermarket, as discussed later—then the difficulty in practice reduces to the reliable measurement of future payoffs to shareholders. Extant research mostly uses short- and long-term analysts' consensus earnings forecasts to proxy for market expectations, but with the reliance on these forecasts two problems arise.

First, analyst forecasts tend to be overly optimistic (e.g., Dugar and Nathan (1995), Francis and Philbrick (1993), McNichols and O'Brien (1997)) which leads to upward biased ICC estimates. This bias in analyst forecasts might even be substantial enough to render an otherwise valid approach unreliable. For example, Easton and Monahan (2005, p. 501) investigate seven different ICC estimates and conclude that "for the entire cross-section of firms these proxies are unreliable", but also show that when "analysts' forecast accuracy is high" some of them become valid. In a similar vein, Guay et al. (2011) show that the association between ICC proxies and future realised returns is weak, but improves substantially once analyst forecast errors are controlled for and, recently, Mohanram and Gode (2013) confirm that removing predictable analyst forecast errors significantly improves the validity of ICC estimates. Second, young, small and financially distressed firms are rarely covered by security analysts, as reported in Diether et al. (2002), Hong et al. (2000), La Porta (1996) and validated by only 22 percent of IPOs in my sample enjoying immediate analyst coverage.[86] This coverage bias limits the ICC methodology to-

[86] Analyst coverage is assumed if at least one analyst provides an earnings forecast for the first fiscal year ending following the IPO.

wards large and well-established firms and, thus, makes it less of an alternative for traditional RFB proxies which only require a firm to have a sufficiently long trading history in order to be estimated.

Given these analyst-associated biases, recent advancements in ICC research recommend mechanical earnings forecasts—either from longitudinal or cross-sectional earnings models—to gauge future payoffs to shareholders. While longitudinal models impose strong data requirements (once again limiting applicability), cross-sectional models remain usable even though firms lack a long time-series of earnings realisations; that is, because factor loadings in the cross-sectional models are estimated via pooled regressions for all firms in the population (say, for the entire Compustat universe), only current firm-specific information for the independent variables is needed to predict future earnings of the sample firms. Three prominent cross-sectional forecasting models are proposed in seminal work by HVZ (2012) and LM (2014) whose applicability to (1) predict earnings and (2) derive implied cost of capital for newly listed firms is tested as part of my main analyses.

4.2.2.1 The HVZ model

HVZ (2012) forecast earnings from contemporary accounting data (viz. total assets, dividends, earnings and accruals) and find that their model-based earnings predictions are superior to analyst forecasts in respect to coverage, forecast bias and earnings response coefficients (ERC). They further show that ICC estimates calculated from model-based earnings forecasts are more reliable proxies for CoE than the ones based on analyst forecasts. The HVZ model builds on previous cross-sectional profitability models (see Fama and French (2000), Fama and French (2006), Hou and Robinson (2006)) and is specified as:

$$E_{i,t+\tau} = \alpha_0 + \alpha_1 A_{i,t} + \alpha_2 D_{i,t} + \alpha_3 DD_{i,t} + \alpha_4 E_{i,t} + \alpha_5 NegE_{i,t} + \alpha_6 AC_{i,t} + \varepsilon \quad (4.5)$$

where $E_{i,t+\tau}$ is earnings in year t+τ (τ = 1 to 5) of firm *i*; $A_{i,t}$ is total assets; $D_{i,t}$ is dividends; $DD_{i,t}$ is an indicator variable for dividend paying firms; $E_{i,t}$ is earnings; $NegE_{i,t}$

is an indicator variable for loss firms; and $AC_{i,t}$ is working capital accruals which I set to zero if missing.

4.2.2.2 The LM models

LM (2014) propose two parsimonious alternatives to the HVZ model in that their earnings persistence (EP) model only requires current earnings data to predict future earnings, and their residual income (RI) model predicts earnings based on current earnings, equity book value and total accruals data.[87] They find that both models outperform the HVZ model in terms of forecast accuracy, forecast bias, ERC, and greater construct validity of the model-based ICC estimates (i.e., greater correlation with future returns and common risk factors). Their earnings persistence (EP) model, specified in (4.6), includes an indicator variable for loss firms ($NegE_{i,t}$) and accounts for different persistence of profits and losses in the form of an interaction term ($NegE_{i,t} * E_{i,t}$).

$$E_{i,t+\tau} = \beta_0 + \beta_1 NegE_{i,t} + \beta_2 E_{i,t} + \beta_3 (NegE * E)_{i,t} + \varepsilon \quad (4.6)$$

Their second model derives from the well-known residual income (RI) model (formalised in a series of influential papers by Ohlson (1995), Feltham and Ohlson (1995, 1996)) and is shown in (4.7).

$$E_{i,t+\tau} = \chi_0 + \chi_1 NegE_{i,t} + \chi_2 E_{i,t} + \chi_3 (NegE * E)_{i,t} + \chi_4 B_{i,t} + \chi_5 TACC_{i,t} + \varepsilon \quad (4.7)$$

where $B_{i,t}$ is equity book value. $TACC_{i,t}$ is total accruals of firm *i* as described in Richardson et al. (2005) and set to zero if missing.

4.2.2.3 The HVZ/EP model

Prior evidence shows that financial losses are less persistent than profits (Hayn (1995), Li (2011), Resutek (2011)). Given that IPO firms tend to show greater propensity to report negative earnings than established firms (e.g., 29 percent of IPOs in my sample are loss firms in year *t* compared to only 20 percent for the Compustat population), it might prove

[87] A discussion of their paper is provided in Feng (2014).

valuable to combine the HZV and EP model into one comprehensive forecasting solution. Thus, I amend the HZV model by an interaction term—see Eq. (4.8)—to account for different earnings persistency ($NegE_{i,t} * E_{i,t}$).

$$E_{i,t+\tau} = \psi_0 + \psi_1 A_{i,t} + \psi_2 D_{i,t} + \psi_3 DD_{i,t} + \psi_4 E_{i,t} + \psi_5 NegE_{i,t} + \psi_6 AC_{i,t} + \psi_7 (NegE * E)_{i,t} + \varepsilon \tag{4.8}$$

4.2.2.4 The RW model

Bradshaw et al. (2012) show that random walk time-series forecasts of earnings compete reasonably well with analyst forecasts over varying forecasting horizons and that these findings are particularly true for younger and smaller companies. Thus, as a naïve benchmark for the cross-sectional earnings forecasting models in this study, I include a random walk (RW) model in the main analysis as shown in Eq. (4.9).

$$E_{i,t+\tau} = E_{i,t} + \varepsilon \tag{4.9}$$

4.3 Methodology

4.3.1 Data and Sample Selection

I use a sample of all NYSE, Amex and Nasdaq listed firms on the Compustat fundamentals annual files up to 2015 to estimate the earnings models. For the HVZ (EP) model 205,484 (208,070) firm-year observations with required data are available from 1950 to 2015. Data coverage for the RI model begins in 1960 with 196,303 firm-year observations available up to 2015. Given that each model is estimated over the past ten years (as described below), the first factor loadings are attainable in 1959 for the HVZ, EP and HVZ/EP model, and in 1969 for the RI model.

My sample of newly listed firms consists of all initial public offerings in the Securities Data Company (SDC) database, excluding unit IPOs, closed-end funds, real estate investment trusts (REITs), American Depositary Receipts and Shares (ADRs & ADSs), non-

U.S. listed firms, and all companies for which pre-listing accounting information is missing on Compustat. Also, I only include IPOs for which the difference between the IPO date (from SDC) and the fiscal year-end (FYE) date (from Compustat) immediately preceding the IPO is at least 90 days and at most 365 days which avoids look ahead bias and reliance on dated fiscal information when forecasting IPO earnings. I sort the remaining 2,485 firms by the difference of days between IPO date and FYE date and confirm for the top and bottom two percent of this sorting that the fiscal year end information used in this study was available in the IPO prospectus (i.e., cross-validating SEC S-1 fillings, provided via EDGAR from 1996 onwards, with Compustat data).[88] Finally, only those IPOs are maintained for which all relevant market (CRSP) and accounting data (Compustat) is available to predict one- to three-year ahead earnings for all four models. The final sample consists of 1,657 IPOs (see Table 4.2).

Table 4.2: IPO Sample Development

Description	**No. of IPOs**
All US Public & Private Common Stock IPOs from 1962 - 2015 listed on SDC Platinum	9,522
Excluding Unit IPOs, REITs, ADRs, ADSs	-364
Excluding IPOs not listed on "U.S Public" market place	-863
Excluding IPOs not to be found on Compustat	-4,452
Intersection SDC & Compustat	**3,843**
Excluding IPOs with missing pre-IPO data on Compustat	-661
Including only IPOs for which difference between IPO date (SDC) and FYE date (Compustat) is greater than or equal to 90 days	-686
Including only IPOs for which difference between IPO date (SDC) and FYE date (Compustat) is smaller than or equal to 365 days	-11
Subtotal (Cross-Validation with EDGAR)	**2,485**
Excluding IPOs for which required market (CRSP) and accounting information (Compustat) is missing	-828
Final Sample (Intersection SDC, Compustat & CRSP)	**1,657**

The table describes IPO sample development in this study. The final sample contains 1,657 IPOs for which all relevant accounting and market data is available to forecast one- to three-year ahead earnings for all four cross-sectional earnings forecasting models employed (HVZ, EP, RI, HVZ/EP).

[88] This analysis also showed that cut-offs below 90 days include accounting information on Compustat which is not available in the IPO prospectus; conversely, cut-offs above 365 days include outdated information.

4.3.2 Earnings Forecasts

Forecasting earnings from cross-sectional models is a two-step procedure: first, factor loadings for each model are attained (using all firms listed on Compustat); second, factor loadings are multiplied by IPOs accounting information; hence, a newly listed firm must only have current financial information available on the independent variables of the respective forecasting model to obtain an estimate of its future earnings. More specifically, in the first step I estimate the HVZ, EP, RI, HVZ/EP model using previous ten years of data from the whole Compustat universe as of the end of each year *t*, which includes all firms with fiscal year ending (FYE) from *June t* to *May t+1*.[89] Consistent with HVZ (2012), all models are estimated in absolute dollar values—and not on per share basis as in LM (2014)—given that the number of shares outstanding reported on Compustat deviates in some instances quite markedly from information reported in the IPO prospectus; accounting information, on the other hand, is reported reliably.[90]

In the second step, for each IPO which occurs between *Jan–Dec t+1* accounting information immediately preceding the IPO date is multiplied by the "right" factor loadings from step one to compute one- to three-year ahead earnings forecasts (four and five-year ahead earnings forecasts available upon request). The only financials available to investors is from the IPO prospectus which generally covers previous 3-years of financial data preceding the year of the IPO. For 1,526 (92%) of the IPOs, the latest available information pertains to fiscal year *t (June t* to *May t+1)* and, therefore, is multiplied by estimated model coefficients for year *t.*[91] For instance, if a firm is initially listed in calendar

89 To mitigate the impact of outliers, all level variables in the respective model are winsorized at the 1st and 99th percentile in each year.

90 I draw randomly ten firms from the IPO sample and cross-checked total assets, net income, book value of equity and shares outstanding figures from the IPO prospectus with Compustat information. Shares outstanding deviated in two from the ten cases, but all other variables have been equal.

91 For 78 (5%) IPOs, the year of the IPO matches the fiscal year end. Such cases arise because IPOs are observed in calendar time (*Jan–Dec t+1*), but fiscal year ends are aligned from *Jun t – May t+1*. Thus, IPOs very late in *t+1* (say, *Dec*) may also have fiscal year end *t+1*, because the fiscal year end information reported and available to investors might be of, say, *Jun t+1*. Conversely, for 53 (3%) IPOs there is a "two-year" difference between the IPO year and the fiscal end year. This can happen for IPOs occurring very early in calendar year *t+1* (say *Jan*), but information reported and available to investors is of *t-2* (say, *May t-1*). To ease readability and given that it holds true for almost all IPOs, we refer from hereafter to *t* as the fiscal year end information available before the IPO, irrespective of as to whether *t+1* or *t-1* fiscal year end information is used (of course, in the former case we multiply by *t+1* and in the latter with *t-1* model coefficients).

year 2006 (*t+1*), the earnings forecasts are in effect based on pre-IPO financial information as of fiscal year 2005, and the year of the first earnings forecast (*t+1)* coincides with the year of the IPO (i.e., 2006). Figure 4.1 illustrates the two-step procedure in obtaining IPO earnings forecasts.

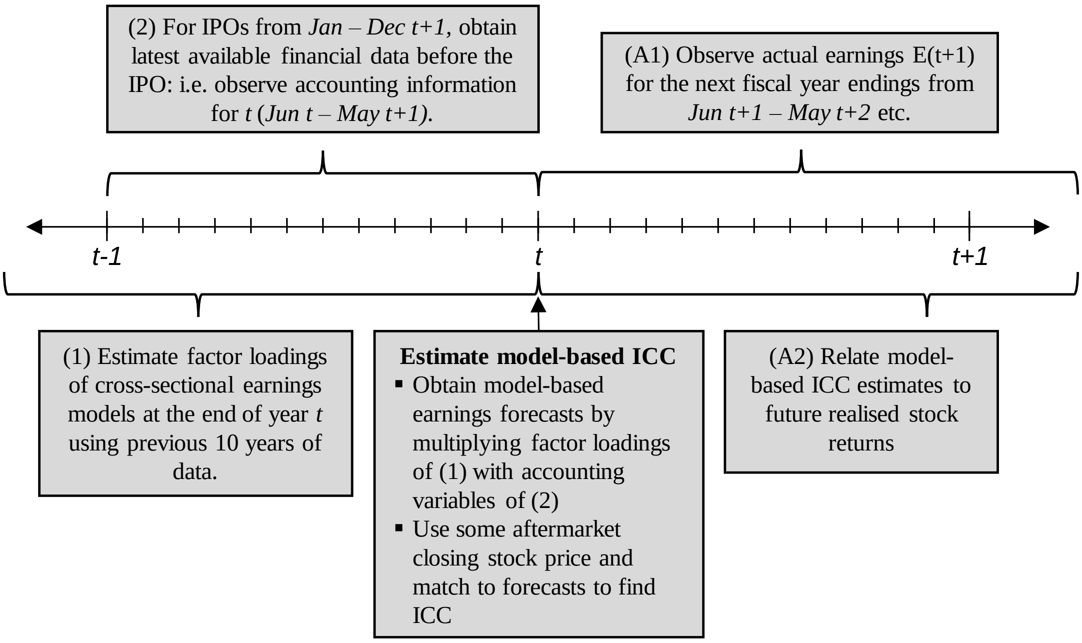


Figure 4.1: Timeline of Earnings Forecasts and ICC estimates

This figure illustrates the two-step procedure in obtaining earnings forecasts for IPOs; outlines how ICC proxies are estimated and indicates data requirements for the two main analyses (A1) and (A2).

To find the "right" factor loadings for the different forecasting horizons (factor loadings for 1-year ahead forecasts are different from 2-year ahead forecasts etc.), the earnings models are estimated on different sets of previous data. That is, pooled cross-sectional regressions for 1-year ahead earnings in year *t* uses data from year *t-11* to *t-2*; regressions for 2-year ahead earnings use data from *t-12 to t-3* and so forth. The two-year gap between *t* and the end of the data period for one-year ahead forecasts (and three-year gap for two-year ahead forecasts etc.) is required to avoid look ahead bias. In the case of one-year ahead forecasts, the first "gap year" is needed because the models regress current accounting data (*t*) on lead earnings (*t+1*). As such, current data must be of *t-1* for lead earnings to be of *t*. The second "gap year" avoids that firms going public very early in calendar year *t+1* use coefficients which are based on earnings that have not yet been reported for fiscal year *t*. As an example, if an IPO occurs in calendar year 2006 *(t+1)* and thus pertains to fiscal year *t* = 2005, then data from 1994-2003 (*t-11 to t-2*) is used to estimate factor loadings that are then applied to calculate earnings for year 2006 (*t+1*); data from 1993-2002 (*t-12 to t-3*) is used to estimate factor loadings that are relevant to calculate earnings in year 2007 (*t+2*) and so forth.

4.3.3 Implied Cost of Capital Estimates

4.3.3.1 Market value of equity

The validity of ICC measures not only hinge upon reliable future earnings estimates, but also on efficient market prices. While market efficiency seems a reasonable assumption in large sample ICC studies, the efficiency of the IPO pricing process remains a puzzling phenomenon (e.g., Cornelli and Goldreich (2003), Lowry and Schwert (2004)). Therefore, the decision which share price is used to obtain an IPO's equity valuation is crucial and the following peculiarities of the IPO pricing process need to be considered. First, the phenomenon of IPO *underpricing* causes offer prices to be on average much lower than first day closing prices; however, IPO firms tend to underperform in the long-run, indicating *overpricing* in the immediate aftermarket (e.g., Ritter (1991), Ritter and Welch (2002)). Second, underwriter price stabilisation activities can severely impact stock prices in the days immediately following the offering (e.g., Hanley (1993), Ruud (1993)). Third, lock-up expirations—occurring mostly 180 days after initial listing—lead to significant price drops of about two percent (Brav and Gompers (2000, 2003), Field and Hanka (2001)).

To increase the likelihood that the share prices used to estimate IPOs implied cost of capital are a fair and efficient reflection of IPOs equity value, I obtain market values at seven different points in time. First, I use the aftermarket closing prices on the 21st day of trading as reported on CRSP (P_1) or, if not available, closing prices at the end of the first month of trading (Lowry et al., 2010). Second, I apply closing prices at the end of sixth month of trading (P_6) given the material impact of lock-up expirations on prices (Brav and Gompers, 2000, 2003). Third, I use closing prices at the end of the ninth (P_9) and eighteenth (P_{18}) month of trading, based on findings in Ecker (2014) that "positive abnormal returns disappear after the first nine post-IPO months" and "negative abnormal returns persist for 18 months" (p. 907). Fourth, I include closing prices at three (P_3), twelve (P_{12}) and twenty-four (P_{24}) months of trading to ensure meaningful comparison over time.

4.3.3.2 ICC methodology

Consistent with HVZ (2012) and LM (2014), I use a composite ICC estimate (rCOMP) in my main analyses which is the average of the four ICC measures suggested by Easton (2004): rPE, rPEG, rMPEG and rAEGM. These estimates are all based on the abnormal earnings growth model and only require one- and two-year ahead earnings forecasts; however, their estimations differ with respect to assumption made about dividend payments, short-term and perpetual abnormal earnings growth. The simplest ICC estimate (rPE) assumes no dividend payout nor any short- or perpetual growth in abnormal earnings. Conversely, rAEGM entertains all available assumptions, with rPEG and rMPEG lying between those two "extreme" cases— see Eq. (4.10) to (4.13). As earnings are forecasted in absolute dollar values, market values of equity (closing price times number of shares outstanding) instead of share prices are used in the calculations. Number of shares outstanding is obtained from CMM (CRSP/Compustat Merged database).

$$r_{PE} = \left(\frac{MV}{E_{t+1}}\right)^{-1} \tag{4.10}$$

$$r_{PEG} = \sqrt{\frac{E_{t+2} - E_{t+1}}{MV}} \tag{4.11}$$

where E_{t+1} (E_{t+2}) equals one (two) year ahead earnings forecasts (in absolute dollar values) from the HVZ, EP, RI and HVZ/EP model; MV equals equity market value based on either closing price at the end of first month of trading (MV1); or closing price at the end of the third (MV3), sixth (MV6), ninth (MV9), twelfth (MV12), eighteenth (MV18), twenty-fourth (MV24) month of trading; whenever MV3, MV6, MV9 or MV12 is applied, earnings forecast for the following year are used to avoid reliance on realised earnings; that is, E_{t+1} is replaced by E_{t+2}; and E_{t+2} becomes E_{t+3}. In case MV18 or MV24 is applied, earnings forecasts for the following two years are used; that is E_{t+1} (E_{t+2}) is replaced by E_{t+3} (E_{t+4}).

$$r_{MPEG} = A + \sqrt{A^2 + \frac{E_{t+2} - E_{t+1}}{MV}} \tag{4.12}$$

where $A = D_{t+1}/2MV$; D_{t+1} equals dividends in *t+1* which are forecasted assuming dividend payout ratio at *t* for firms with positive earnings, or using dividends at *t* divided by six percent of total assets at *t* as a proxy for dividend payout ratio for firms with current negative earnings (Gebhardt et al., 2001).

$$r_{AEGM} = A + \sqrt{A^2 + \frac{E_{t+1}}{MV}\left(\frac{E_{t+2} - E_{t+1}}{E_{t+1}} - G_{AEG}\right)} \tag{4.13}$$

where $A = \frac{1}{2}\left(G_{AEG} + \frac{D_{t+1}}{MV}\right)$; G_{AEG} = perpetual growth rate in abnormal earnings set to current risk-free rate minus three percent (Gode and Mohanram, 2003).[92]

4.3.4 Performance Measures

4.3.4.1 Forecast bias and accuracy

I apply forecast bias and forecast accuracy as the main performance measures to examine the quality of forecasted IPO earnings. Forecast bias is defined as the difference between actual/realised earnings and predicted earnings: a negative bias denotes overpredicted earnings and optimistic forecasts. To allow for a meaningful comparison between IPOs, forecast bias is scaled by market value of equity at the end of first month of trading (MV1). For each model, forecast biases are winsorized at the 1st and 99th percentile since very few observations severely distort mean figures. Forecast accuracy is the absolute value of forecast biases (*alias* absolute forecast error).

92 The current risk free rate is calculated by first taking the average of the one-month Treasury bill rate over rolling past 12-month windows and then annualising these averages by compounding over 12 months (Barth et al., 2013). Treasury bill rates are from Kenneth French website.

4.3.4.2 Earnings response coefficient

Discussed in seminal work by Ball and Brown (1968) and Beaver (1968), earnings response coefficients (ERC) are direct measures of market-expectations about future earnings in that they measure stock price reaction to one unit of unexpected earnings, with higher ERCs indicating a higher "quality of the earnings expectation model employed" (Kothari, 2001, p. 117). Given that the measurement of "announcement" ERCs depend on the availability of market-adjusted returns (which are unavailable for IPOs given their lack of trading history), I estimate "annual" ERCs instead. That is, every year buy-and-hold returns (i.e., monthly compounded CRSP total returns) over the next one-, two- and three years are regressed on unexpected earnings (i.e., forecast bias) over the same period (HVZ, 2012).[93] I require at least 15 IPOs for every annual regression and standardise unexpected earnings to have mean zero and unit variance in each year.

4.3.4.3 Fama-MacBeth regressions

The second part of my analysis investigates the degree to which forecasted earnings can be used to derive valid ICC estimates for IPOs. I follow Guay et al. (2011) and evaluate the performance of each estimate by running Fama-MacBeth regressions of one-year ahead buy-and-hold returns on the composite ICC (rCOMP), where I require at least 15 IPOs in each fiscal year and set all ICC estimates to a range of 0 and 0.50 as it is unlikely that investors would expect returns below zero or above fifty percent (Barth et al., 2013).[94,95] Depending on which market value is used to calculate the ICC measures, it is assumed that shares are either bought at the end of the first, third, sixth, ninth, 12th, 18th or 24th month of trading and hold for the subsequent 12 months (i.e., monthly compounding of total returns reported on CRSP); these future realised returns are then regressed on the corresponding ICC estimates. For instance, ICC measures calculated from market

93 In the one-year example, shares are bought at the end of first month of trading and then hold for 12 months; these buy-and-hold returns (BHR) are then regressed on each models forecast bias for E(t+1). In the two-year example, BHR are based on a 24-months holding period and regressed on the sum of forecast biases E(t+1) and E(t+2). The three-year example follows accordingly.

94 Easton and Monahan (2005) extend Guya et al. and control for cash flow and discount rate news in their regressions. However, this approach requires forecast revisions which are not attainable from cross-sectional models (LM, 2014, footnote 8).

95 Another approach to analyse the construct validity of ICC measures is to examine their association with common risk factors (e.g., market beta, leverage, B/M, volatility, size). Easton and Monahan (2016), however, argue that this approach is illogical, given that the ICC approach assumes that these very risk factors are either unknown or cannot be reliably measured.

prices at the end of the sixth month of trading (i.e., MV6) are regressed on future 12months buy-and-hold returns, where stocks are also bought at the end of the sixth month of trading. The benchmark coefficient of theses univariate regressions is one and statistically significant (insignificant) different from zero (one).

4.4 Analysis and Results

4.4.1 IPO Sample Summary Statistics

Table 4.3 reports descriptive statistics for the IPO sample. Consistent with prior evidence (e.g., Fama and French (2004)), IPO activity peaked in the 1990s with a total of 885 new issues (53 percent), followed by 435 IPOs (26 percent) in the 1980s; the average IPOs market cap after one month of trading is 485mUSD, and about 18 percent of IPOs (293) relate to the Business Service industry, followed by a combined 14 percent of IPOs belonging to the Computer (120) and Electronic Equipment (112) sector, respectively. For only 362 IPOs (22 percent) immediate IBES coverage is available, which highlights the limitations of the common analyst-based ICC approach for new joiners in capital markets and underpins the importance to elaborate further on "analyst-independent" methods to provide newly listed firms with reliable CoE estimates.[96]

As shown in Table 4.4, the average IPO firm in my sample is about 12-17 times smaller than the average Compustat firm in respect to book value of equity (B_{IPO}: 66.3mUSD vs. B_{COM}: 994.3), total assets (A_{IPO}: 402.9 vs. A_{COM}: 3,118.2), earnings (E_{IPO}: 8.2 vs. E_{COM}: 103.4), dividends (D_{IPO}: 4.1 vs. D_{COM}: 36.5), working capital accruals (AC_{IPO}: -4.2 vs. AC_{COM}: -50.0) and total accruals ($TACC_{IPO}$: 1.7 vs. $TACC_{COM}$: 28.3). IPOs are also less likely to pay dividends (20 percent), but more likely to report negative earnings (29 percent) in the fiscal year before their initial listing vis-à-vis the average firm on Compustat (dividend payers: 53 percent; loss firms: 20 percent). These statistics show that my sample is based on smaller, younger and less followed firms in equity markets which corroborates the objective of this paper to examine the applicability of the HVZ and LM models for those types of firms.

[96] Analyst coverage is assumed if at least one IBES forecast is provided for E(t+1).

Table 4.3: Distribution of IPO Sample

Panel A: Distribution by calendar year, IBES coverage and market capitalisation

Year	Freq.	IBES	IBES (%)	Mkt. Cap	Year	Freq.	IBES	IBES (%)	Mkt. Cap
1972	1	-	-	43.1	1993	118	23	19.5	24,869.6
1973	1	-	-	82.1	1994	104	27	26.0	15,655.3
1974	-	-	-	-	1995	101	24	23.8	38,815.7
1975	-	-	-	-	1996	149	28	18.8	74,644.5
1976	3	-	-	203.9	1997	85	26	30.6	41,683.8
1977	-	-	-	-	1998	55	22	40.0	27,310.4
1978	1	-	-	44.9	1999	69	15	21.7	144,161.4
1979	1	-	-	25.7	2000	51	14	27.5	51,191.5
1980	13	-	-	1,163.6	2001	12	1	8.3	31,582.5
1981	40	1	2.5	3,021.0	2002	14	1	7.1	34,824.2
1982	11	3	27.3	948.3	2003	16	5	31.3	7,289.0
1983	94	13	13.8	10,179.9	2004	40	16	40.0	32,346.3
1984	37	3	8.1	2,095.4	2005	38	16	42.1	25,024.3
1985	30	1	3.3	3,311.6	2006	33	11	33.3	23,674.9
1986	94	6	6.4	11,488.5	2007	32	12	37.5	59,242.5
1987	66	13	19.7	8,604.5	2008	5	2	40.0	6,959.1
1988	25	6	24.0	2,899.8	2009	7	3	42.9	8,466.9
1989	25	10	40.0	6,015.4	2010	18	5	27.8	12,423.6
1990	25	7	28.0	2,359.3	2011	18	8	44.4	25,907.0
1991	79	13	16.5	13,072.4	2012	25	5	20.0	17,378.1
1992	100	18	18.0	18,349.2	2013	21	4	19.0	17,079.9
					Total	**1,657**	**362**	**21.8**	**804,438.9**

Panel B: Distribution by 48 Fama and French (1997) industries

Industry Name	Freq.	Freq. (%)	Industry Name	Freq.	Freq. (%)
Agriculture	6	0.36	Aircraft	7	0.42
Food Products	17	1.03	Shipbuilding, Railroad Eq.	3	0.18
Candy & Soda	1	0.06	Defence	1	0.06
Alcoholic Beverages	5	0.30	Precious Metals	2	0.12
Tobacco Products	1	0.06	Non-metallic Mining	1	0.06
Recreational Products	16	0.97	Coal	1	0.06
Entertainment	26	1.57	Petroleum and Natural Gas	36	2.17
Printing and Publishing	8	0.48	Utilities	30	1.81
Consumer Goods	27	1.63	Telecommunications	48	2.90
Apparel	23	1.39	Personal Services	26	1.57
Healthcare	38	2.29	Business Services	293	17.68
Medical Equipment	87	5.25	Computers	120	7.24
Pharmaceutical Products	103	6.22	Electronic Equipment	112	6.76
Chemicals	10	0.60	Measuring and Control Eq.	46	2.78
Rubber and Plastic Products	9	0.54	Business Supplies	6	0.36
Textiles	7	0.42	Shipping Containers	6	0.36
Construction Materials	19	1.15	Transportation	42	2.53
Construction	7	0.42	Wholesale	51	3.08
Steel Works	18	1.09	Retail	99	5.97
Fabricated Products	2	0.12	Restaurants, Hotels, Motels	38	2.29
Machinery	50	3.02	Banking	51	3.08
Electrical Equipment	17	1.03	Insurance	36	2.17
Miscellaneous	7	0.42	Real Estate	9	0.54
Automobiles and Trucks	13	0.78	Trading	76	4.59
			Total	**1,657**	**100.00**

Panel A reports number of IPOs by calendar year. Market Cap. (mUSD) is calculated as shares outstanding x closing price either at the end of the 21st day of trading or—if not available—end of first month after a firm going-pubic. IBES reports number of IPOs for which at least one analyst provides E(t+1) forecasts on IBES. Panel B reports the number of IPOs by 48 Fama and French (1997) industries.

Table 4.4: IPO Sample vs. Compustat Population Summary Statistics

Panel A: IPO Sample Summary Statistics of Variables Used in Forecasting Earnings

Variable	Mean	1%	25%	Median	75%	99%	STD	Obs.
B(t)	66.31	-173.16	-0.28	4.78	17.97	1,048.30	661.68	1,657
A(t)	402.94	1.25	10.17	28.54	105.94	5,399.50	3,212.43	1,657
E(t)	8.24	-40.83	-0.58	1.39	5.02	165.22	60.48	1,657
NegE(t)	0.29	0.00	0.00	0.00	1.00	1.00	0.45	1,657
NegE*E(t)	-2.53	-40.83	-0.58	0.00	0.00	0.00	9.80	1,657
D(t)	4.06	0.00	0.00	0.00	0.00	95.63	33.99	1,657
DD(t)	0.20	0.00	0.00	0.00	0.00	1.00	0.40	1,657
AC(t)	-4.15	-85.51	0.00	0.00	0.00	19.53	49.14	1,657
TACC(t)	1.68	-61.06	0.00	0.00	0.00	86.78	90.66	1,657

Panel B: Population Summary Statistics of Variables Used in Estimating Earnings models

Variable	Mean	1%	25%	Median	75%	99%	STD	Obs.
B(t)	994.33	-116.82	18.45	78.41	354.58	17,035.00	5,932.06	209,443
A(t)	3,118.16	2.13	43.02	203.24	1,073.61	56,732.00	14,505.20	221,850
E(t)	103.41	-123.89	0.48	5.30	34.44	2,211.97	494.67	208,070
NegE(t)	0.20	0.00	0.00	0.00	0.00	1.00	0.40	208,070
NegE*E(t)	-6.29	-123.89	0.00	0.00	0.00	0.00	48.61	208,070
D(t)	36.54	0.00	0.00	0.20	7.00	731.90	181.50	219,800
DD(t)	0.53	0.00	0.00	1.00	1.00	1.00	0.50	219,800
AC(t)	-49.90	-1,192.00	-7.06	0.00	0.00	88.10	357.80	250,876
TACC(t)	28.31	-451.50	0.00	0.00	3.73	937.70	958.76	250,876

*Panel A reports summary statics for the variables used to forecast earnings from the HVZ, EP, RI and HVZ/EP model. Panel B reports summary statistics for the same variables for the entire Compustat population from 1950 – 2015. All level variables are winsorised at 1^{st} and 99^{th} percentile. All values in mUSD, except for NegE, DD and NegE*E).*

4.4.2 Model-Based Earnings Forecasts

4.4.2.1 Coefficient estimates of the cross-sectional models

Required data for the HVZ and EP model are available from 1950 to 2015 and data coverage for the RI model begins in 1960. As each model is estimated over the past ten years, the first estimation period for the HVZ, EP and HVZ/EP models end in 1959 and for the RI model in 1969. Average coefficients for each model are reported in Table 4.5 and are widely consistent with estimates in HVZ (2012) and LM (2014).

Table 4.5: Coefficient Estimates from the HVZ, EP & RI Models

Panel A: The HVZ model, 1959 - 2015

	Intercept	A(t)	D(t)	DD(t)	E(t)	NegE(t)	AC(t)	Adj. R^2
E(t+1)	2.899	0.003	0.243	0.527	0.848	1.789	-0.033	0.89
	(4.17)	*(4.46)*	*(3.93)*	*(2.15)*	*(29.05)*	*(2.41)*	*(-2.49)*	
E(t+2)	5.234	0.005	0.336	1.295	0.797	2.831	-0.046	0.83
	(6.33)	*(5.91)*	*(4.61)*	*(3.18)*	*(21.71)*	*(2.44)*	*(-2.78)*	
E(t+3)	8.011	0.008	0.405	0.836	0.776	2.730	-0.046	0.79
	(7.90)	*(6.95)*	*(4.68)*	*(3.37)*	*(18.39)*	*(1.89)*	*(-2.47)*	

Panel B: The EP model, 1959 - 2015

	Intercept	NegE(t)	E(t)	NegE*E(t)	Adj. R^2
E(t+1)	4.351	-8.321	1.014	-1.196	0.88
	(6.79)	*(-4.54)*	*(100.64)*	*(-7.85)*	
E(t+2)	7.871	-12.004	1.051	-1.951	0.82
	(10.04)	*(-5.33)*	*(78.58)*	*(-8.54)*	
E(t+3)	11.004	-13.817	1.105	-2.433	0.78
	(11.88)	*(-5.82)*	*(69.95)*	*(-8.59)*	

Panel C: The RI model, 1969 - 2015

	Intercept	NegE(t)	E(t)	NegE*E(t)	B(t)	TACC(t)	Adj. R^2
E(t+1)	5.088	-9.635	0.987	-1.000	0.004	-0.007	0.88
	(7.03)	*(-5.00)*	*(71.62)*	*(-7.89)*	*(2.73)*	*(-1.28)*	
E(t+2)	8.920	-13.483	1.013	-1.630	0.007	-0.011	0.82
	(9.69)	*(-5.71)*	*(50.73)*	*(-8.43)*	*(3.03)*	*(-1.10)*	
E(t+3)	12.139	-15.086	1.060	-2.102	0.009	-0.020	0.78
	(11.69)	*(-5.83)*	*(44.38)*	*(-8.16)*	*(3.21)*	*(-1.53)*	

Panel D: The HVZ/EP model, 1959 - 2015

	Intercept	A(t)	D(t)	DD(t)	E(t)	NegE(t)	AC(t)	NegE*E(t)	Adj. R^2
E(t+1)	3.116	0.003	0.224	-0.119	0.867	-7.429	-0.028	-0.860	0.89
	(4.48)	*(3.87)*	*(3.58)*	*(-1.73)*	*(29.33)*	*(-4.15)*	*(-2.15)*	*(-5.59)*	
E(t+2)	5.574	0.005	0.310	0.316	0.825	-10.322	-0.040	-1.396	0.83
	(6.73)	*(5.25)*	*(4.20)*	*(2.67)*	*(22.31)*	*(4.87)*	*(-2.41)*	*(-6.18)*	
E(t+3)	8.347	0.007	0.375	-0.182	0.808	-11.823	-0.038	-1.663	0.80
	(8.29)	*(6.30)*	*(4.27)*	*(-2.86)*	*(19.04)*	*(-5.26)*	*(-2.09)*	*(-6.17)*	

Each model is estimated annually using previous ten-years of data from 1959-2015 for the HVZ, EP, HVZ/EP and 1969-2015 for the RI models. For each model, average mean coefficients and time-series average t-statistic based on robust standard errors (in parentheses) are reported.

4.4.2.2 Forecast bias and accuracy

Table 4.6, Panel A reports forecast bias for each of the four cross-sectional earnings forecasting models (including the random walk model) and pairwise comparisons between each of them are provided in Panel B. Figure 4.2 summarises this information graphically.

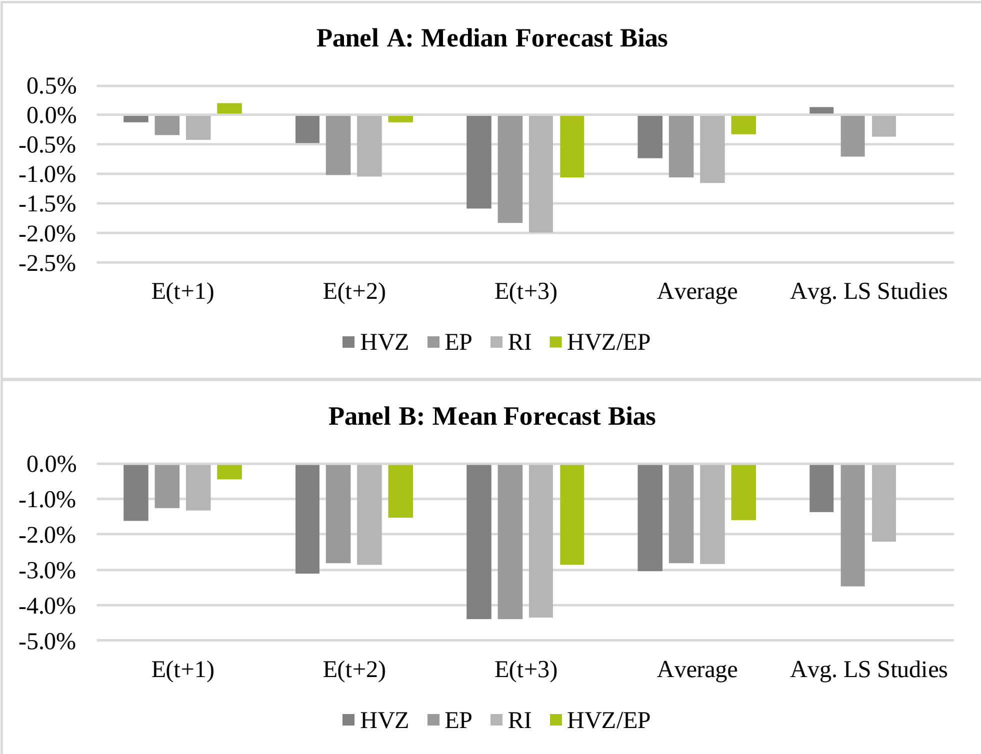


Figure 4.2: Forecast Bias in Percent of Market Value

Panel A (Panel B) shows median (mean) forecast bias in percent of market value for each of the four cross-sectional earnings forecasting models for one- to three-year ahead earnings forecasts. For comparison, the average forecast bias in large sample studies (Avg. LS studies) is included in this figure and based on data shown in Table 4.8.

The HVZ, EP and RI models show similar mean forecast biases over the three forecasting horizons of about -1.5 percent of market value for E(t+1); -3.0 and -4.5 percent for E(t+2) and E(t+3), respectively. All biases are significantly negative and increasing over time, which indicates that one-year ahead forecast are less overpredicted than three-year ahead ones. Comparing median figures, HVZ model forecasts are about 1 to 3 times less biased than EP and RI model predictions [HVZ: E(t+1) -0.13; E(t+2) -0.47; E(t+3) -1.59] with all median differences being highly significant. The amended HVZ/EP model produces the least mean [E(t+1): -0.45; E(t+2): -1.53; E(t+3): -2.85] and median biases [E(t+1): 0.20; E(t+2): -0.13; E(t+3): -1.00] among all models. This is illustrated by highly significant differences between HVZ/EP biases and the remaining three models (only the three-year ahead median difference between HVZ-HVZ/EP is insignificant; z-sore: -1.85). The RW model consistently underpredicts IPO earnings for E(t+1) to E(t+3) with an average forecast bias of approximately 2 percent of market value.

Table 4.7 reports forecast accuracy figures and shows that the HVZ, EP, RI and HVZ/EP model generate similarly accurate predictions for one- to three-year ahead earnings (see also Figure 4.3). Forecasting precision is decreasing over time for all four models as mean (median) accuracy is increasing from a low of about 3.7 (2.1) percent for E(t+1) to a high of 9.1 (5.4) percent for E(t+3) forecasts. Earnings predictions from the RW model are as accurate as those from the cross-sectional models. This is consistent with findings in Bradshaw et al. (2012) who show that in particular for younger and smaller firms naïve earnings forecasts compete reasonable well with analyst forecasts over varying forecasting horizons. Taken together, all five models perform equally well in regard to the accuracy of their respective forecasts, but the HVZ/EP model generates less biased earnings predictions than the HVZ, EP and RI models.

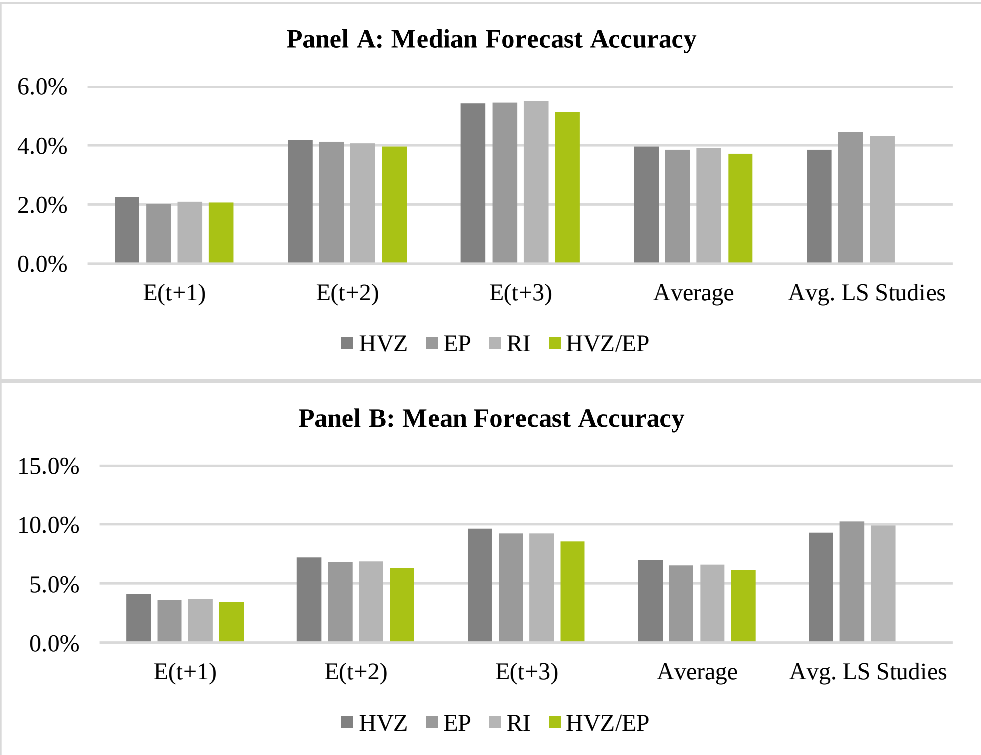


Figure 4.3: Forecast Accuracy in Percent of Market Value

Panel A (Panel B) shows median (mean) forecast accuracy in percent of market value for each of the four cross-sectional earnings forecasting models for one- to three-year ahead earnings forecasts. For comparison, average forecast accuracy in large sample studies (Avg. LS studies) is included in this figure and based on data shown in Table 4.8.

Table 4.6: Forecast *Bias* of the Forecasting Models for the IPO Sample

Panel A: Time-Series Averages

	E(t+1)		E(t+2)		E(t+3)	
	Mean	Median	Mean	Median	Mean	Median
HVZ	-0.0163	-0.0013	-0.0311	-0.0047	-0.0440	-0.0159
	(-9.98)	*(5.91)*	*(-11.20)*	*(-7.55)*	*(-12.01)*	*(-10.16)*
EP	-0.0127	-0.0033	-0.0282	-0.0102	-0.0439	-0.0182
	(-9.17)	*(7.13)*	*(-11.22)*	*(-9.21)*	*(-12.83)*	*(-11.58)*
RI	-0.0132	-0.0041	-0.0287	-0.0104	-0.0435	-0.0199
	(9.21)	*(7.44)*	*(-11.30)*	*(-9.37)*	*(-12.76)*	*(-11.62)*
HVZ/EP	-0.0045	0.0020	-0.0153	-0.0013	-0.0285	-0.0105
	(-3.54)	*(-0.14)*	*(-6.56)*	*(-3.75)*	*(-8.93)*	*(-7.21)*
RW	0.0134	0.0134	0.0198	0.0220	0.0216	0.0227
	(12.95)	*(17.01)*	*(10.52)*	*(13.94)*	*(8.32)*	*(11.64)*

Panel B: Pairwise Comparison

	E(t+1)		E(t+2)		E(t+3)	
	Mean	Median	Mean	Median	Mean	Median
HVZ-EP	-0.0036	0.0021	-0.0028	0.0055	-0.0001	0.0024
	(-2.97)	*(4.06)*	*(-1.55)*	*(6.12)*	*(-0.07)*	*(7.53)*
HVZ-RI	-0.0031	0.0029	-0.0023	0.0057	-0.0005	0.0041
	(-2.46)	*(4.16)*	*(-1.25)*	*(6.16)*	*(-0.23)*	*(7.34)*
HVZ-HVZ/EP	-0.0118	-0.0033	-0.0158	-0.0035	-0.0155	-0.0054
	(-10.58)	*(-5.05)*	*(-9.36)*	*(-3.82)*	*(-8.55)*	*(-1.85)*
HVZ-RW	-0.0297	-0.0147	-0.0509	-0.0267	-0.0656	-0.0386
	(-24.79)	*(-33.18)*	*(-26.73)*	*(-34.45)*	*(-28.15)*	*(-34.92)*
EP-RI	0.0005	0.0008	0.0005	0.0002	-0.0004	0.0017
	(4.13)	*(2.69)*	*(2.78)*	*(3.42)*	*(-0.58)*	*(2.42)*
EP-HVZ/EP	-0.0082	-0.0054	-0.0130	-0.0089	-0.0154	-0.0077
	(-13.96)	*(-21.18)*	*(-14.19)*	*(-22.53)*	*(-15)*	*(-22.48)*
EP-RW	-0.0261	-0.0168	-0.0480	-0.0322	-0.0655	-0.0409
	(-27.16)	*(-31.85)*	*(-29.12)*	*(-33.18)*	*(-29.24)*	*(-33.66)*
RI-HVZ/EP	-0.0087	-0.0062	-0.0134	-0.0092	-0.0150	-0.0094
	(-13.98)	*(-21.58)*	*(-14.37)*	*(-22.9)*	*(-12.69)*	*(-22.45)*
RI-RW	-0.0267	-0.0176	-0.0485	-0.0324	-0.0651	-0.0427
	(-26.06)	*(-31.46)*	*(-28.71)*	*(-32.97)*	*(-28.23)*	*(-33.43)*
HVZ/EP-RW	-0.0179	-0.0114	-0.0351	-0.0232	-0.0501	-0.0332
	(-21.21)	*(-28.31)*	*(-23.50)*	*(-30.66)*	*(-25.51)*	*(-32.16)*

Panel A reports time-series averages of the mean and median forecast biases for the cross-sectional earnings forecasting models and the random walk model (RW). In parentheses time-series t-statistics for mean and median (Wilcoxon signed-rank test) are reported. Panel B reports pairwise comparisons (differences) between the models with t-statistics for mean (paired t-test) and median (Wilcoxon signed rank sum test) differences reported in parentheses. Results are based on 1,657 IPOs.

Table 4.7: Forecast *Accuracy* of the Forecasting Models for the IPO Sample

Panel A: Time-Series Averages

	E(t+1)		E(t+2)		E(t+3)	
	Mean	Median	Mean	Median	Mean	Median
HVZ	0.0408	0.0227	0.0725	0.0418	0.0966	0.0542
	(30.14)	*(35.26)*	*(32.09)*	*(35.26)*	*(32.22)*	*(35.26)*
EP	0.0364	0.0204	0.0681	0.0412	0.0924	0.0547
	(32.95)	*(35.26)*	*(33.97)*	*(35.26)*	*(33.27)*	*(35.26)*
RI	0.0372	0.0212	0.0687	0.0408	0.0927	0.0552
	(32.12)	*(35.26)*	*(33.89)*	*(35.26)*	*(33.60)*	*(35.26)*
HVZ/EP	0.0341	0.0208	0.0633	0.0398	0.0860	0.0514
	(35.49)	*(35.26)*	*(35.69)*	*(35.26)*	*(34.46)*	*(35.26)*
RW	0.0306	0.0205	0.0555	0.0379	0.0740	0.0490
	(39.02)	*(35.26)*	*(40.01)*	*(35.26)*	*(38.40)*	*(35.26)*

Panel B: Pairwise Comparison

	E(t+1)		E(t+2)		E(t+3)	
	Mean	Median	Mean	Median	Mean	Median
HVZ-EP	0.0044	0.0024	0.0044	0.0006	0.0042	-0.0004
	(4.79)	*(5.48)*	*(3.19)*	*(2.25)*	*(2.51)*	*(1.74)*
HVZ-RI	0.0036	0.0016	0.0038	0.0010	0.0039	-0.0010
	(3.84)	*(4.75)*	*(2.64)*	*(2.05)*	*(2.25)*	*(1.84)*
HVZ-HVZ/EP	0.0067	0.0019	0.0091	0.0021	0.0106	0.0028
	(8.11)	*(9.69)*	*(7.54)*	*(8.01)*	*(7.36)*	*(7.04)*
HVZ-RW	0.0102	0.0022	0.0169	0.0039	0.0226	0.0052
	(8.73)	*(2.04)*	*(8.91)*	*(3.21)*	*(9.56)*	*(5.22)*
EP-RI	-0.0008	-0.0008	-0.0006	0.0004	-0.0002	-0.0005
	(-6.20)	*(-4.63)*	*(-3.54)*	*(-2.42)*	*(-0.74)*	*(-1.37)*
EP-HVZ/EP	0.0023	-0.0005	0.0048	0.0015	0.0065	0.0032
	(4.35)	*(0.94)*	*(5.87)*	*(4.27)*	*(6.73)*	*(6.07)*
EP-RW	0.0058	-0.0002	0.0126	0.0033	0.0185	0.0057
	(6.07)	*(0.50)*	*(7.37)*	*(2.54)*	*(8.07)*	*(4.56)*
RI-HVZ/EP	0.0031	0.0003	0.0054	0.0011	0.0067	0.0038
	(5.50)	*(1.51)*	*(6.40)*	*(4.32)*	*(6.67)*	*(5.80)*
RI-RW	0.0066	0.0006	0.0132	0.0029	0.0187	0.0062
	(6.52)	*(1.07)*	*(7.57)*	*(2.74)*	*(8.13)*	*(4.55)*
HVZ/EP-RW	0.0035	0.0003	0.0078	0.0018	0.0120	0.0024
	(4.69)	*(-1.40)*	*(5.68)*	*(0.08)*	*(6.36)*	*(1.95)*

Panel A reports time-series averages of the mean and median forecast accuracy for the cross-sectional earnings forecasting models and the random walk model (RW). In parentheses time-series t-statistics for mean and median (Wilcoxon signed-rank test) are reported. Panel B reports pairwise comparisons (differences) between the models with t-statistics for mean (paired t-test) and median (Wilcoxon signed rank sum test) differences reported in parentheses. Results are based on 1,657 IPOs.

4.4.2.3 Comparison with large sample studies

Table 4.8 compares forecasts biases (Panel A) and accuracies (Panel B) of this study's IPO earnings forecasts with large sample study results of HVZ (2012, Table 3) and LM (2014, Table 2 & 3). Given that my forecast biases are winsorised at the 1st and 99th percentile, which contrasts with HVZ and LM, I focus the discussion on median figures.

HVZ forecast bias for the IPO sample increases from a low of -0.13 percent for E(t+1) to a high of -1.69 percent for E(t+3), and lies between figures reported in HVZ (2012) and LM (2014): the former reports bias as low as zero percent and the latter as high as -4.3 percent. EP model predictions for the IPO sample are about twice as biased than in LM (2014) for E(t+1) and E(t+2), but similar for E(t+3); results for the RI model show a similar pattern, but differences in biases are of larger magnitude. On the other hand, IPO forecast biases for the HVZ/EP model are highly comparable with figures in HVZ (2012) and LM (2014). In respect to forecast accuracy it is shown that IPO earnings forecasts for E(t+1) generated from HVZ, EP and RI models are about 30 percent more accurate than in large sample studies; EP and RI based forecasts of E(t+2) and E(t+3) remain more accurate than in LM (2014), whereas for the same forecasting horizons HVZ predictions become less accurate if compared against HVZ (2012). Similar in magnitude to HVZ, EP and RI stand-alone, HVZ/EP forecast accuracy compares equally well to large sample study results over all forecasting horizons.

4.4.2.4 Earnings response coefficients

Table 4.9 provides time-series averages of "annual" ERCs for the IPO earnings forecasts (Panel A) and pairwise comparisons across the models are provided in Panel B. Figure 4.4 summarises this information graphically.

Earnings response coefficients for one-, two- and three-year ahead earnings predictions from the HVZ (0.09, 0.25, 0.37) and HVZ/EP (0.09, 0.27, 0.40) model are somewhat higher than those from the EP (0.08, 0.23, 0.36) and RI (0.08, 0.23, 0.35) model; however, these differences are insignificant (Panel B). Notably, ERCs are approximately 60 percent lower than figures reported in HVZ (2012, Table 3), implying reduced value relevance of earnings for IPO firms vis-à-vis more mature and stable firms. This is consistent with prior evidence showing that it is not only financial (e.g., earnings), but also non-financial information that explains IPO valuations (e.g., Bartov et al. (2002), Klein (1996)). Annual ERCs from the RW model (0.08, 0.26, 0.49) are similar to those from HVZ, EP, RI and HVZ/EP for E(t+1) and E(t+2), but significantly higher for E(t+3), which is similar to findings in Bradshaw et al. (2012) who compare analyst-based earnings forecasts with naïve RW predictions.

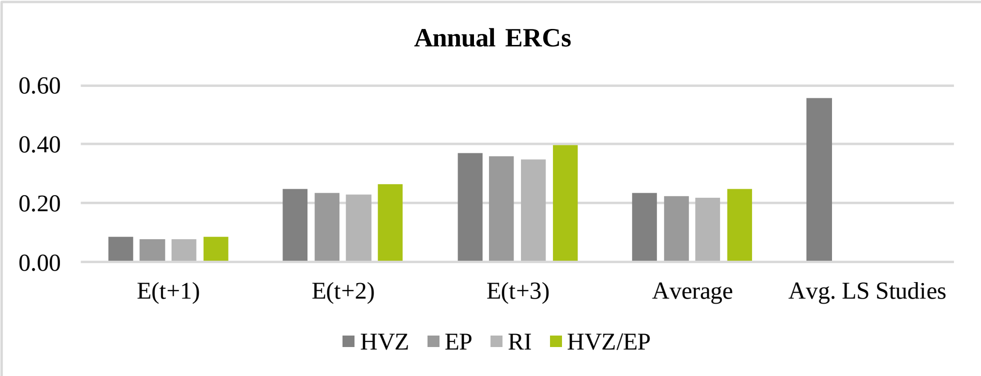


Figure 4.4: Earnings Response Coefficients

The figure shows time-series averages of "annual" earnings response coefficients (ERCs) for each of the four cross-sectional earnings forecasting models. The average large study ERC (Avg. LS studies) reported in this figure is based on HVZ (2012, Table 3).

To summarise, results suggest a slight outperformance of the HVZ/EP model in forecasting IPO earnings (mostly because of its lower forecast biases and higher ERC coefficients), followed by HVZ in second place, and EP & RI in joint third. This ranking differs from findings in LM (2014)—where it is shown that EP and RI earnings predictions perform better than HVZ forecasts—which indicates that model performance varies with firm characteristics: while earnings in large sample studies (and, thus, for more established firms) can be reliably predicted by means of parsimonious forecasting models (i.e., EP and RI), newly listed firms require more comprehensive forecasting solutions (even beyond HVZ) to minimise bias and maximise accuracy of forecasted earnings, with the HVZ/EP model being one such solution.

Table 4.8: Comparison of Forecast Bias & Accuracy with Large Sample Studies

Panel A: Comparison of forecast biases

	E(t+1)		E(t+2)		E(t+3)		
	Mean	Median	Mean	Median	Mean	Median	Obs.
HVZ/EP	-0.0045	0.0020	-0.0153	-0.0013	-0.0285	-0.0105	1,657
HVZ							
IPOs	-0.0163	-0.0013	-0.0311	-0.0047	-0.0440	-0.0159	1,657
HVZ (2012)	-0.0209	0.0031	-0.0167	0.0010	-0.0109	-0.0001	87,825
LM (2014)	-0.0560	-0.0130	-0.0920	-0.0270	-0.1330	-0.0430	119,653
EP							
IPOs	-0.0127	-0.0033	-0.0282	-0.0102	-0.0439	-0.0182	1,657
LM (2014)	-0.0200	0.0000	-0.0320	-0.0060	-0.0520	-0.0150	119,653
RI							
IPOs	-0.0132	-0.0041	-0.0287	-0.0104	-0.0435	-0.0199	1,657
LM (2014)	-0.0130	0.0020	-0.0190	-0.0030	-0.0340	-0.0100	119,653
RW							
IPOs	0.0134	0.0134	0.0198	0.0220	0.0216	0.0227	1,657
LM (2014)	0.0070	0.0080	0.0270	0.0150	0.0400	0.0220	119,653

Panel B: Comparison of forecast accuracies

	E(t+1)		E(t+2)		E(t+3)		
	Mean	Median	Mean	Median	Mean	Median	Obs.
HVZ/EP	0.0341	0.0208	0.0633	0.0398	0.0860	0.0514	1,657
HVZ							
IPOs	0.0408	0.0227	0.0725	0.0418	0.0966	0.0542	1,657
HVZ (2012)	0.0837	0.0302	0.0938	0.0400	0.1031	0.0458	87,825
LM (2014)	0.1010	0.0340	0.1510	0.0570	0.2030	0.0790	119,653
EP							
IPOs	0.0364	0.0204	0.0681	0.0412	0.0924	0.0547	1,657
LM (2014)	0.0730	0.0280	0.1010	0.0450	0.1330	0.0610	119,653
RI							
IPOs	0.0372	0.0212	0.0687	0.0408	0.0927	0.0552	1,657
LM (2014)	0.0730	0.0270	0.0990	0.0440	0.1260	0.0590	119,653
RW							
IPOs	0.0306	0.0205	0.0555	0.0379	0.0740	0.0490	1,657
LM (2014)	0.0880	0.0280	0.1020	0.0430	0.1280	0.0560	119,653

The table compares forecast biases (Panel A) and accuracies (Panel B) for this studies IPO earnings forecasts with large sample results from HVZ (2012, Table 3) and LM (2014, Table 2). HVZ (2012, Table 3) is based on observations for which both model- and analyst-based forecasts are available; thus, sample size varies by horizon: E(t+1): 99,100; E(t+2): 89,454; E(t+3): 74,922 and I report the average between the three figures.

Table 4.9: ERC of the Forecasting Models for the IPO Sample

Panel A: Time-Series Averages

	E(t+1)		E(t+2)		E(t+3)	
	ERC	R-squared	ERC	R-squared	ERC	R-squared
HVZ	0.0865	0.0093	0.2485	0.0168	0.3714	0.0199
	(2.95)		*(4.38)*		*(4.48)*	
EP	0.0788	0.0094	0.2341	0.0145	0.3585	0.0195
	(2.92)		*(4.35)*		*(4.88)*	
RI	0.0782	0.0091	0.2288	0.0136	0.3498	0.0182
	(3.03)		*(4.39)*		*(4.92)*	
HVZ/EP	0.0855	0.0119	0.2650	0.0201	0.3972	0.0260
	(3.06)		*(5.09)*		*(5.16)*	
RW	0.0714	0.0090	0.2635	0.0232	0.4905	0.0362
	(2.11)		*(4.06)*		*(4.87)*	

Panel B: Pairwise Comparison

	E(t+1)		E(t+2)		E(t+3)	
	ERC	R-squared	ERC	R-squared	ERC	R-squared
HVZ-EP	0.0077	-0.0001	0.0145	0.0023	0.0129	0.0004
	(0.19)		*(0.37)*		*(0.33)*	
HVZ-RI	0.0083	0.0002	0.0198	0.0032	0.0217	0.0017
	(0.21)		*(0.52)*		*(0.56)*	
HVZ-HVZ/EP	0.0010	-0.0026	-0.0164	-0.0033	-0.0258	-0.0061
	(0.02)		*(-0.43)*		*(-0.63)*	
HVZ-RW	0.0151	0.0003	-0.0150	-0.0064	-0.1190	-0.0163
	(0.34)		*(-0.33)*		*(-2.44)*	
EP-RI	0.0005	0.0003	0.0053	0.0009	0.0088	0.0013
	(0.01)		*(0.14)*		*(0.25)*	
EP-HVZ/EP	-0.0068	-0.0025	-0.0309	-0.0056	-0.0387	-0.0065
	(-0.17)		*(-0.83)*		*(-1.02)*	
EP-RW	0.0074	0.0004	-0.0294	-0.0087	-0.1319	-0.0167
	(0.17)		*(-0.68)*		*(-2.84)*	
RI-HVZ/EP	-0.0073	-0.0028	-0.0362	-0.0065	-0.0475	-0.0078
	(-0.19)		*(-0.99)*		*(-1.27)*	
RI-RW	0.0069	0.0001	-0.0347	-0.0096	-0.1407	-0.0180
	(0.16)		*(-0.81)*		*(-3.06)*	
HVZ/EP-RW	0.0142	0.0029	0.0015	-0.0031	-0.0932	-0.0102
	(0.32)		*(0.03)*		*(-1.95)*	

Panel A reports time-series averages of annual earnings response coefficients (ERC) for the cross-sectional earnings forecasting models and the random walk model (RW). In parentheses Fama-MacBeth t-statistics. Panel B reports pairwise comparisons between the models, with t-statistics for ERC differences reported in parentheses (paired t-test). Results are based on 1,590; 1,587 and 1,493 observations for E(t+1), E(t+2) and E(t+3), respectively.

4.4.3 Model-Based Implied Cost of Capital Estimates

4.4.3.1 Descriptive statistics and correlations

Table 4.10 reports summary statistics for realised and expected returns. One-year ahead buy-and-hold returns (BHR) are strongly positively skewed over all seven points in time. Average BHR of 17.3 percent are realised if stocks are bought at MV1, but subsequently decline to about 12 percent if purchased at MV3 and MV6. Beginning with MV9 realised returns are monotonically increasing from 14.9 percent to a high of 20.0 percent for MV24. Median returns display a similar pattern over time and range from negative 2.2 percent to positive 5.5 percent. These findings are incompatible with large sample studies—for example, Guay et al. (2011) report realised one-year-ahead returns of about 13 percent for both mean and median—, but can be explained by the return distribution of my IPO sample: some IPOs perform exceptionally well—181 (149) [149] yield annual returns of above 100 percent if bought at MV1 (MV6) [MV12]—while the median IPO hardly creates any return for investors.

Expected returns for the composite ICC are highly consistent across all four models and over time. The average (median) IPO in my sample has an expected rate of return of 10.9 (8.6) percent, which is identical to average analyst-based ICC estimates in Liu et al. (2014) for a sample of approximately 800 IPOs.[97] Figures also compare well with both model- and analyst-based ICC estimates in large sample studies: HVZ (2012) document a mean (median) of 14.9 (9.2) percent for their model-based ICC composite and Botosan et al. (2011) report an average (median) of about 10.4 (10.1) percent for their analyst-based composites.[98] The somewhat lower median levels in this study can be explained by the fact that many IPOs in my sample (53 percent) pertain to the 1990s, a decade which marks the beginning of declining median ICC estimates to about 6 percent per year (HVZ, 2012, Table 4).

[97] Appendix 4.1 provides mean and median expected returns for each of the four ICC measures (rPE, rPEG, rMPEG, rAEGM) from which rCOMP is calculated.

[98] HVZ (2012) composite is formed of rGLS, rCT, rOJN (similar to rAEGM), rMPEG and rGOR (similar to rPE); Botosan et al. (2011) uses two composites: rHL consists of rCT, rGLS, rMPEG, and rOJN (similar to rAEGM), and rDKL is the average of rCT, rGLS and rGM (similar to rMPEG).

Table 4.10: Pooled Average Future Realised Returns and Implied Cost of Capital

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
rREAL														
Actual	0.1725	0.0088	0.1214	-0.0193	0.1219	-0.0222	0.1493	0.0126	0.1706	0.0270	0.1820	0.0256	0.2008	0.0545
Winsor 1%	0.1546	0.0088	0.1124	-0.0193	0.1056	-0.0222	0.1327	0.0126	0.1426	0.0270	0.1654	0.0256	0.1858	0.0545
Winsor 2%	0.1491	0.0088	0.1029	-0.0193	0.0929	-0.0222	0.1176	0.0126	0.1275	0.0270	0.1549	0.0256	0.1713	0.0545
rCOMP														
HVZ	0.1113	0.0919	0.1029	0.0827	0.1032	0.0816	0.1045	0.0808	0.1069	0.0830	0.1094	0.0810	0.1092	0.0811
EP	0.1147	0.1003	0.1059	0.0883	0.1060	0.0859	0.1076	0.0845	0.1104	0.0852	0.1216	0.0955	0.1219	0.0933
RI	0.1148	0.0998	0.1051	0.0858	0.1051	0.0850	0.1067	0.0832	0.1094	0.0846	0.1214	0.0952	0.1215	0.0939
HVZ/EP	0.1017	0.0860	0.0987	0.0810	0.0987	0.0797	0.0999	0.0787	0.1023	0.0797	0.1111	0.0848	0.1111	0.0852

The table reports pooled mean and median future realised returns and expected returns for the composite ICC measure rCOMP. HVZ, EP, RI and HVZ/EP denote the model which is used to estimate rCOMP. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Realised return observations from MV1 to MV24: 1590/1591/1592/1592/1591/1563/1504. Expected return observations for all models from MV1 to MV24:1600/1594/1594/1592/1592/1592/1591.

Table 4.11 shows temporal averages of year-by-year Spearman and Pearson correlations between realised and expected returns, including the number of years (out of 29 for MV1-MV18 and 28 for MV24) for which annual coefficients are significantly positive/negative. Average Spearman correlations for the HVZ, EP, RI and HVZ/EP model centre around zero up to MV9 (i.e., half of the sixteen significant correlations are positive and half of them are negative), but become consistently positive from MV12 onwards (Panel A). Across all models, HVZ-based estimates show the highest association with realised returns for MV12 (0.07), MV18 (0.06) and MV24 (0.05), followed by HVZ/EP with an average correlation of 0.04 over the same forecasting horizons. The lowest correlations are observed for EP- and RI-based estimates, with a mean association of about 0.03. Pearson correlations (Panel B) lead to qualitatively similar conclusions; however, the linear relationships between realised and expected returns tend to be considerably stronger than the monotonic ones above; that is, the average correlation between HVZ-based ICC estimates and actual returns from MV12 to MV24 is about 0.09, and approximately 0.08 for EP, RI and HVZ/EP.

While comparable in magnitude to prior evidence—for instance, Botosan et al. (2011) document an average Spearman correlation of about 0.06—results differ in respect to statistical significance: approximately forty percent of coefficients are significantly positive in large sample studies (e.g., Botosan and Plumlee (2005), Botosan et al. (2011)), while only about ten (twenty) percent of Spearman (Pearson) coefficients are positive for my IPO sample (averaging correlations over all models and forecasting horizons beyond MV9). This study's correlation coefficients are, on average, based on 60 observations per year; Botosan and Plumlee (2005) on the other hand rely on approximately 1,100 observations (12,400 firm-years from 1983-1993) and Botosan et al. (2011) infer results from about 700 observations per year (14,521 firm-years from 1984–2004). Thus, low sample size possibly diminishes the statistical power of these tests and explains the comparatively low number of significant firm-year correlations for my IPO sample.

Table 4.11: Year-by-Year Correlations

Panel A: Spearman correlations

	MV1	MV3	MV6	MV9	MV12	MV18	MV24
	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*
rCOMP							
HVZ	-0.0172	0.0148	0.0214	-0.0026	0.0686	0.0578	0.0471
	(2/1)	*(1/1)*	*(3/1)*	*(2/0)*	*(3/0)*	*(3/1)*	*(3/1)*
EP	0.0082	0.0194	0.0003	-0.0331	0.0361	0.0271	0.0226
	(4/1)	*(3/2)*	*(4/2)*	*(2/1)*	*(2/0)*	*(3/1)*	*(3/0)*
RI	-0.0001	0.0145	-0.0018	-0.0382	0.0273	0.0283	0.0229
	(3/1)	*(3/2)*	*(4/2)*	*(2/1)*	*(3/0)*	*(3/2)*	*(2/0)*
HVZ/EP	-0.0078	0.0065	0.0011	-0.0439	0.0203	0.0474	0.0458
	(3/1)	*(2/1)*	*(3/2)*	*(3/0)*	*(4/0)*	*(2/1)*	*(4/0)*

Panel B: Pearson correlations

	MV1	MV3	MV6	MV9	MV12	MV18	MV24
	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*	Correl. *(+/-)*
rCOMP							
HVZ	-0.0592	-0.0106	0.0441	0.0052	0.0994	0.0787	0.0832
	(0/2)	*(2/1)*	*(3/2)*	*(2/2)*	*(5/0)*	*(5/1)*	*(7/0)*
EP	-0.0313	-0.0014	0.0098	-0.0417	0.0799	0.0712	0.0748
	(1/0)	*(3/1)*	*(3/3)*	*(2/2)*	*(4/0)*	*(6/1)*	*(4/1)*
RI	-0.0321	-0.0045	0.0107	-0.0425	0.0786	0.0768	0.0723
	(1/0)	*(3/2)*	*(3/2)*	*(2/1)*	*(5/0)*	*(6/1)*	*(4/1)*
HVZ/EP	-0.0479	-0.0185	-0.0065	-0.0571	0.0581	0.0754	0.0957
	(1/2)	*(3/1)*	*(2/3)*	*(1/2)*	*(5/0)*	*(5/0)*	*(5/0)*

Panel A (Panel B) reports the average value of year-by-year Spearman (Pearson) correlations between realised and expected returns. In parentheses the number of years (out of 29 for MV1-MV18 and 28 for MV24) for which annual coefficients are significantly positive/negative at the 0.05 level. HVZ, EP, RI and HVZ/EP denote the model which is used to predict earnings. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Results based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24.

4.4.3.2 Empirical results

Table 4.12 provides Fama-MacBeth coefficients from regressing one-year ahead buy-and-hold returns on the model-based ICC measures. The benchmark for each regression is a coefficient of one which indicates that realised and expected returns are on average equal. LM (2014) report statistically significant non-zero coefficients of 0.31 (3.42) for HVZ-based ICC estimates, 0.57 (2.59) for EP and 0.65 (2.81) for RI (Fama-MacBeth t-statistic in parentheses).[99]

FMB coefficients for HVZ, EP, RI and HVZ/EP are all negative and insignificant for MV1, somewhat mixed (7 negative; 5 positive) for MV3 to MV9, but significantly positive for forecasting horizons beyond. HVZ/EP-based ICC estimates display the strongest relationship with future realised returns in that coefficients are consistently close to one (MV12: 0.94; MV18: 1.04; MV24: 0.96), statistically non-zero (t-statistic MV12-MV24: 1.81; 2.52; 2;69) and highly insignificantly different from one (t-statistics MV12-MV24: 0.10). In contrast, HVZ, EP and RI-based estimates tend to be on average too low for MV12 (as indicated by coefficients above one: 1.31; 1.22; 1.12), but too high for MV24 (as illustrated by coefficients below one: 0.87; 0.74; 0.75).[100]

Inefficient market prices in the short-term may explain why results only become significant beyond MV9. Consistent with Miller (1977), investors might diverge on opinions about fundamental values of newly listed shares in the immediate aftermarket. In combination with short-sale constrains—which prohibit pessimistic investors to register their negative outlook—this can result in upward biased aftermarket prices towards the value perceptions of optimistic investors. Only as short-sale constrains are lowered and investor opinions are less divergent, observed market prices converge towards stocks' fundamental values (i.e., market prices become efficient).[101] The effect of this adjustment process is twofold; first, average realised returns are declining until prices reflect fundamental

[99] HVZ (2012) do not carry out firm-level analyses.

[100] Given that scale effects can bias coefficients and confound inferences in regression analyses (e.g., Barth and Clinch (2009), Easton and Sommers (2003)), I re-run the analyses on share-deflated earnings and use closing prices (i.e. P1 to P24) instead of market values (i.e. MV1 to MV24) in the ICC calculations. ICC level estimates remain almost unchanged and FMB coefficients are qualitatively equal to results using undeflated earnings (see Appendix 4.3) which corroborates results.

[101] See also Cornelli and Yilmaz (2015) who model the impact of short-selling constraints on long run equilibrium prices and discuss implications of their model for IPOs.

value (leading to insignificant associations with realised returns); second, only as markets turn efficient, realised and expected returns become positively related. My findings support this proposition: median realised returns are initially decreasing (see Figure 4.5)—which leads to insignificant FMB coefficients up to MV9—but realised returns converge towards expected returns from MV12 onwards (implying and increase in market efficiency) which results in positive and significant FMB coefficients thereafter (see Figure 4.6).

These results are also supported by prior evidence. For instance, Cornelli et al. (2006, p. 1190) find for the vast majority of their sample that (1) retail investors' overoptimism results in first-day IPO prices which are on average 41 percent higher than they would have been in the "absence of sentiment demand" and (2) that prices decline in the first 12 months of trading as "overoptimism gives way to realistic expectations." In a similar vein, Ecker (2014) shows that "positive abnormal returns disappear after the first nine post-IPO months" and "negative abnormal returns persist for 18 months" (p. 907). I interpret these findings as evidence that post-IPO market efficiency increases with time and prices "turn" efficient somewhere between 9 and 18 months of listing.

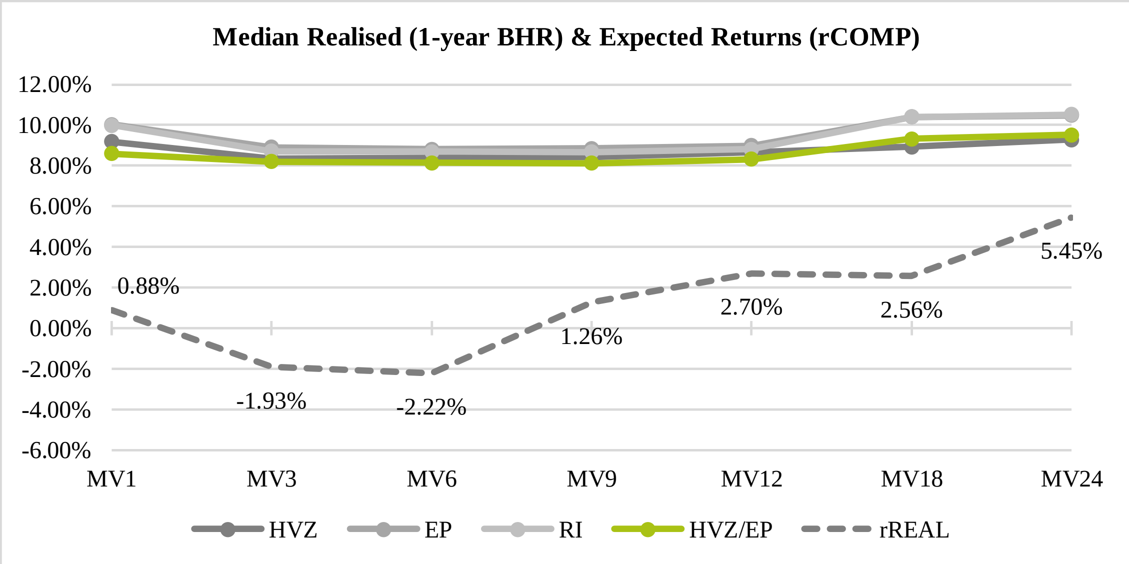
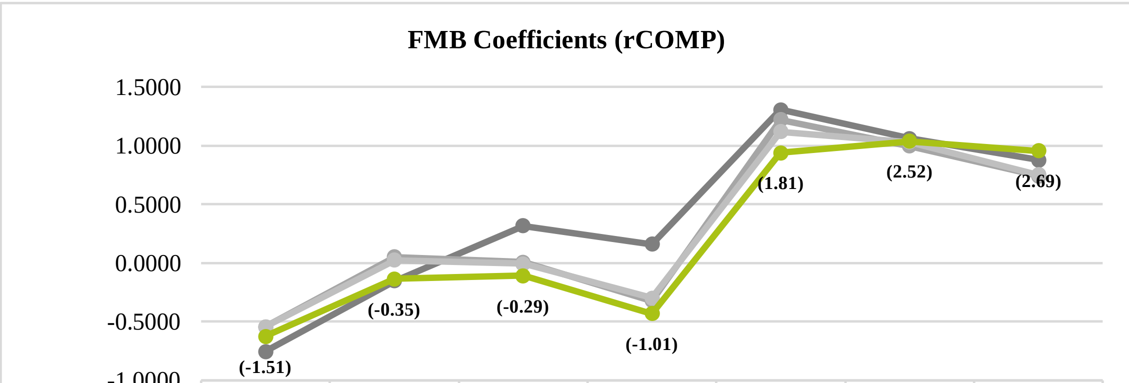


Figure 4.5: Realised & Expected Returns

The figure compares median realised buy-and-hold returns (dotted line) with median expected ICC (rCOMP) returns, calculated from HVZ, EP, RI and HVZ/EP earnings forecasts (solid lines) over time.



	MV1	MV3	MV6	MV9	MV12	MV18	MV24
HVZ	-0.7581	-0.1528	0.3150	0.1597	1.3059	1.0623	0.8789
EP	-0.5459	0.0505	0.0072	-0.3230	1.2236	0.9960	0.7437
RI	-0.5461	0.0248	-0.0064	-0.3012	1.1186	1.0324	0.7539
HVZ/EP	-0.6254	-0.1369	-0.1094	-0.4316	0.9403	1.0383	0.9573

Figure 4.6: Fama-MacBeth Coefficients

The figure shows Fama-MacBeth coefficients from regressing realised returns on expected ICC (rCOMP) returns, calculated from HVZ, EP, RI and HVZ/EP earnings forecasts, over time. HVZ/EP t-statistics are reported in bold parentheses.

Table 4.12: Fama-MacBeth Regressions of Future Annual Returns on ICC estimates

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
rCOMP	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*
HVZ	-0.7581	0.30	-0.1528	0.02	0.3150	0.60	0.1597	0.05	1.3059	0.61	1.0623	1.86	0.8789	0.92
= 0	*(-1.83)**		*(-0.44)*		*(0.76)*		*(0.41)*		*(2.87)****		*(2.59)***		*(2.44)***	
= 1	*(4.24)****		*(3.30)****		*(1.65)*		*(2.15)***		*(0.67)*		*(0.14)*		*(0.33)*	
EP	-0.5459	0.09	0.0505	0.06	0.0072	0.06	-0.3230	0.00	1.2236	0.94	0.9960	1.42	0.7437	0.64
= 0	*(-1.21)*		*(0.13)*		*(0.02)*		*(-0.66)*		*(1.87)**		*(2.34)***		*(2.00)**	
= 1	*(3.43)****		*(2.36)***		*(2.55)***		*(2.70)****		*(0.35)*		*(0.00)*		*(0.69)*	
RI	-0.5461	0.08	0.0248	0.06	-0.0064	0.03	-0.3012	0.00	1.1186	0.87	1.0324	1.57	0.7539	0.63
= 0	*(-1.19)*		*(0.06)*		*(-0.02)*		*(-0.67)*		*(1.80)**		*(2.45)***		*(1.98)**	
= 1	*(3.38)****		*(2.41)***		*(2.56)***		*(2.89)****		*(0.20)*		*(0.01)*		*(0.65)*	
HVZ/EP	-0.6254	0.29	-0.1369	0.02	-0.1094	0.07	-0.4316	0.02	0.9403	0.41	1.0383	1.60	0.9573	0.91
= 0	*(-1.51)*		*(-0.35)*		*(-0.29)*		*(-1.01)*		*(1.81)**		*(2.52)***		*(2.69)****	
= 1	*(3.93)****		*(2.88)****		*(2.92)****		*(3.34)****		*(0.10)*		*(0.10)*		*(0.10)*	

*The table reports Fama-MacBeth regression coefficients of future realised returns on the composite ICC measure rCOMP. In parentheses t-statistics for whether the mean coefficient is equal to zero or one. ***, **, * denotes significance at the one, five, ten percent level or better. HVZ, EP, RI and HVZ/EP denote the model which is used to predict earnings. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Results based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24.*

4.4.3.3 Robustness Tests

To further substantiate the validity of the suggested ICC methodology, I perform several robustness tests. Given above findings, which show that the HVZ/EP model generates more valid ICC estimates vis-à-vis the HVZ, EP and RI models stand-alone, I run those tests for the HVZ/EP model only.

First, it is examined if results are robust to different econometric implementations (see Table 4.13 and Figure 4.7). In the base case—which equals previous results—all ICC estimates are set to a range of zero and 50 percent, and at least 15 IPOs per fiscal year are required, before FMB regressions are estimated. Test I follows data requirements/manipulations of the base case, but operates pooled OLS regressions (controlling for time-fixed effects) instead. Except for MV12, all coefficients are very similar in size and statistical significance to previous results. Data requirements of the base case are undone in Test II in that *actual* future returns are regressed on *actual* ICC estimates, without requiring a minimum number of firms per year. Test coefficients are in the median about 20 percent smaller than under the base case; however, the pattern between the two setups remains highly consistent (Pearson r: 0.93), leaving results qualitatively unchanged. Further, it is examined if results are driven by survivorship bias. The total number of firms in each regression depends on when stocks are assumed to be bought; observations are close to 1,590 for MV1 to MV12, but drop to 1,563 and 1,504 for MV18 and MV24, respectively; whilst small, this indicates survivorship bias in my sample. Keeping only those firms with non-missing data for MV1-MV24 (1,503 observations) and comparing FMB coefficients against previous results (i.e., the base case), it is shown that variation in firm numbers has no material impact on findings (see Test III).

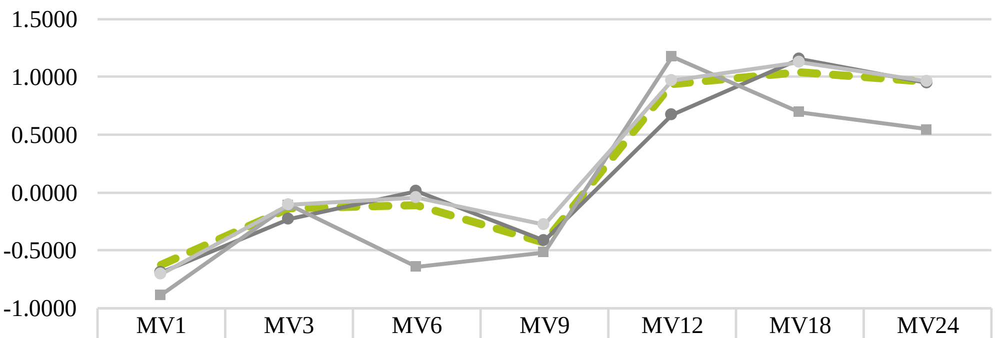
Second, I test if excluding the largest and smallest IPOs (in terms of market value at the end of the first month of trading) from the sample alters results. This is important as IPO size is strongly positively skewed (25.9) and highly leptokurtic (828.2), which might allow extreme cases to bias results. Table 4.14 and Figure 4.8 show FMB coefficients for three different size percentiles: MV[1-99] excludes IPOs with market value below the 1^{st} and above the 99^{th} percentile from the analysis, MV[5-95] excludes firms outside the

$5^{th}/95^{th}$ percentile and MV[10-90] firms outside the $10^{th}/90^{th}$ percentile. Test coefficients between the base case and the different size percentiles are highly correlated (average Pearson r: 0.95), but somewhat lower in terms of magnitude and statistical significance. This indicates slightly upward biased results due to an "abnormally" high representation of very small and large IPOs in my sample.

Third, I analyse if results are robust to different sampling periods. As shown in Table 4.15 and Figure 4.9, FMB coefficients for IPOs pertaining to the 1980s show a very similar pattern to the base case; the 1990s coefficients are slightly higher, and the 2000s are somewhat lower. Only few IPOs in my sample relate to the 1980s (average observations MV1-MV24: 440) and 2000s (234), which partially explains lower statistical significance of the coefficients; in contrast, t-statistics for the 1990s (900) compare well with full sample results. Finally, it is examined if the dot-com bubble (2000/02) and financial crisis (2007/08) distort our findings. To provide coarse indication of these crises' impact on results, I exclude IPOs occurring immediately before and during the collapse of each bubble: fiscal years 1997-2001 (dot-com) and 2004-2007 (financial crisis) are removed from the analysis which—as noted in Section 4.3.2 Earnings Forecasts—include firms going public from 1998-2002 and 2005-2008. FMB coefficients become more conservative under this setting (reduced size and significance) which reveals some upward bias in full sample results due to market inefficiencies surrounding the dot-com and financial crisis.

Taken together, several additional tests support the validity of this study's ICC methodology for newly listed firms. Base case results are somewhat too optimistic due to the impact of very large/small IPOs in my sample as well as distorting effects emanating from the two major financial crises in the early 2000s; however, even after controlling for these confounding factors, the overall construct validity of HVZ/EP-based ICC estimates remain widely intact.

Regression Coefficients (HVZ/EP-rCOMP) for Different Econometric Setups

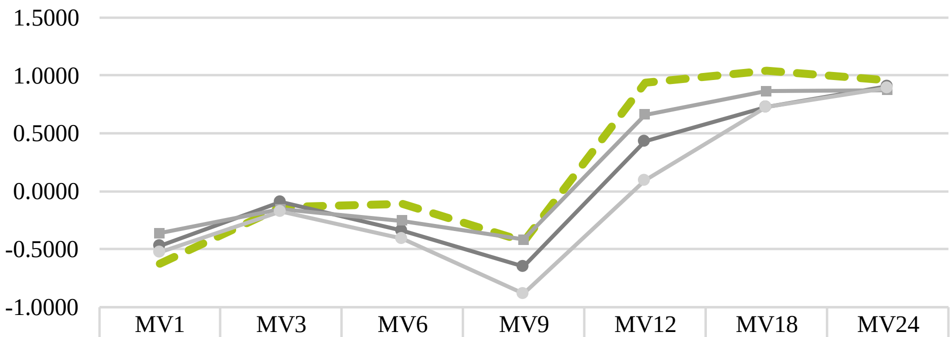


	MV1	MV3	MV6	MV9	MV12	MV18	MV24
Base	-0.6254	-0.1369	-0.1094	-0.4316	0.9403	1.0383	0.9573
Test I	-0.6935	-0.2311	0.0118	-0.4146	0.6728	1.1558	0.9462
Test II	-0.8877	-0.1087	-0.6407	-0.5172	1.1769	0.6951	0.5467
Test III	-0.7041	-0.1033	-0.0461	-0.2768	0.9697	1.1289	0.9590

Figure 4.7: Robustness Tests of HVZ/EP Model for Different Econometric Settings

The figure shows regression coefficients for different econometric settings. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***Test I:*** *Base case, but pooled OLS regressions (controlling for time-fixed effects) are run;* ***Test II:*** *FMB regressions on raw data (i.e., data manipulation and requirements of base case undone);* ***Test III:*** *Base case, but only those firms are kept which have non-missing data for all points in time (MV1-MV24: 1,503 observations).*

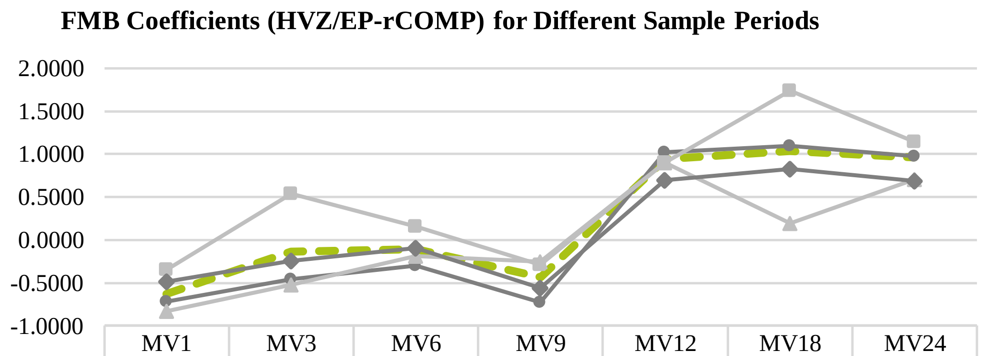
FMB Coefficients (HVZ/EP-rCOMP) for Different IPO Size Percentiles



	MV1	MV3	MV6	MV9	MV12	MV18	MV24
Base	-0.6254	-0.1369	-0.1094	-0.4316	0.9403	1.0383	0.9573
MV [1-99]	-0.4706	-0.0943	-0.3412	-0.6484	0.4307	0.7270	0.9076
MV [5-95]	-0.3644	-0.1566	-0.2558	-0.4195	0.6608	0.8635	0.8722
MV [10-90]	-0.5284	-0.1761	-0.4115	-0.8832	0.0901	0.7298	0.8927

Figure 4.8: Robustness Tests of HVZ/EP Model for Different IPO Size Percentiles

The figure shows FMB regression coefficients for different IPO size percentiles. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***MV[1-99]:*** *Base case, but IPOs with market value below 1st and above 99th percentile are excluded from the analysis; MV[5-95] and MV[10-90] follows accordingly. Percentiles based on market value at the end of the first month of trading.*



	MV1	MV3	MV6	MV9	MV12	MV18	MV24
Base	-0.6254	-0.1369	-0.1094	-0.4316	0.9403	1.0383	0.9573
1980/89	-0.7203	-0.4616	-0.2964	-0.7275	1.0208	1.0981	0.9789
1990/99	-0.3501	0.5373	0.1517	-0.2948	0.8996	1.7389	1.1396
2000/13	-0.8260	-0.5253	-0.1919	-0.2549	0.8961	0.1933	0.7026
Ex.Crises	-0.4850	-0.2410	-0.0910	-0.5608	0.6945	0.8298	0.6909

Figure 4.9: Robustness Tests of HVZ/EP Model for Different Sample Periods

The figure shows FMB regression coefficients of the HVZ/EP model for different sample periods. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP from 1980 - 2013, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***1980/89:*** *Base case, but IPOs pertaining to fiscal years 1980 to 1989 only; 1990/99 and 2000/13 follows accordingly.* ***Ex.Crises:*** *Base case, but IPOs pertaining to fiscal years 1997/2001 and 2004/07 excluded from sample.*

Table 4.13: Robustness Tests of HVZ/EP Model for Different Econometric Settings

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
HVZ/EP-rCOMP	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*
Base	-0.6254	0.29	-0.1369	0.02	-0.1094	0.07	-0.4316	0.02	0.9403	0.41	1.0383	1.60	0.9573	0.91
= 0	(-1.51)		(-0.35)		(-0.29)		(-1.01)		(1.81)*		(2.52)**		(2.69)***	
= 1	(3.93)***		(2.88)***		(2.92)***		(3.34)***		(0.10)		(0.10)		(0.10)	
Test I	-0.6935	6.06	-0.2311	6.50	0.0118	5.39	-0.4146	5.14	0.6728	4.38	1.1558	6.46	0.9462	4.06
= 0	(-2.54)***		(-0.92)		(0.04)		(-1.53)		(2.19)**		(4.71)***		(3.74)***	
= 1	(6.18)***		(4.90)***		(3.73)***		(5.22)***		(1.06)		(0.63)		(0.22)	
Test II	-0.8877	0.00	-0.1087	0.00	-0.6407	0.00	-0.5172	0.02	1.1769	0.10	0.6951	1.09	0.5467	0.28
= 0	(-1.52)		(-0.15)		(-1.46)		(-0.97)		(2.32)**		(1.52)		(1.39)	
= 1	(3.24)***		(1.50)		(3.75)***		(2.85)***		(0.35)		(0.67)		(1.15)	
Test III	-0.7041	0.42	-0.1033	0.01	-0.0461	0.07	-0.2768	0.01	0.9697	0.48	1.1289	1.89	0.9590	0.92
= 0	(-1.61)		(-0.24)		(-0.11)		(-0.56)		(1.76)*		(2.68)***		(2.69)***	
= 1	(3.89)***		(2.60)**		(2.59)**		(2.59)**		(0.00)		(0.30)		(0.10)	

The table reports regression coefficients of the HVZ/EP model for a number of different econometric settings. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***Test I:*** *Base case, but pooled OLS regressions (controlling for time-fixed effects) are run;* ***Test II:*** *FMB regressions on raw data (i.e., data manipulation and requirements of base case undone);* ***Test III:*** *Base case, but only those firms are kept which have non-missing data for all points in time (MV1-MV24: 1,503 observations). In parentheses t-statistics for whether the mean coefficient is equal to zero or one. ***, **, * denotes significance at the one, five, ten percent level or better. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Results based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24, except for Test III (MV1-MV24: 1,503 observations).*

Table 4.14: Robustness Tests of HVZ/EP Model for Different Sample Percentiles

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
HVZ/EP-rCOMP	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*
Base	-0.6254	0.29	-0.1369	0.02	-0.1094	0.07	-0.4316	0.02	0.9403	0.41	1.0383	1.60	0.9573	0.91
= 0	*(-1.51)*		*(-0.35)*		*(-0.29)*		*(-1.01)*		*(1.81)**		*(2.52)***		*(2.69)****	
= 1	*(3.93)****		*(2.88)****		*(2.92)****		*(3.34)****		*(0.10)*		*(0.10)*		*(0.10)*	
MV[1-99]	-0.4706	0.32	-0.0943	0.02	-0.3412	0.00	-0.6484	0.03	0.4307	0.27	0.7270	1.10	0.9076	0.91
= 0	*(-1.02)*		*(-0.25)*		*(-0.92)*		*(-1.56)*		*(1.03)*		*(1.84)**		*(2.50)***	
= 1	*(3.19)****		*(2.85)****		*(3.61)****		*(3.97)****		*(1.36)*		*(0.69)*		*(0.26)*	
MV[5-95]	-0.3644	0.29	-0.1566	0.00	-0.2558	0.00	-0.4195	0.06	0.6608	0.24	0.8635	1.31	0.8722	0.93
= 0	*(-0.95)*		*(-0.39)*		(-0.64)		*(-0.90)*		*(1.47)*		*(2.03)***		*(2.43)***	
= 1	*(3.55)****		*(2.91)****		*(3.15)****		*(3.06)****		*(0.75)*		*(0.32)*		*(0.36)*	
MV[10-90]	-0.5284	0.37	-0.1761	0.00	-0.4115	0.00	-0.8832	0.04	0.0901	0.23	0.7298	1.59	0.8927	0.94
= 0	*(-0.98)*		*(-0.35)*		(-0.79)		*(-1.53)*		*(0.19)*		*(1.57)*		*(2.16)***	
= 1	*(2.83)****		*(2.31)***		*(2.70)****		*(3.26)****		*(1.90)**		*(0.58)*		*(0.26)*	

The table reports FMB regression coefficients of the HVZ/EP model for different sample percentiles. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***MV[1-99]:*** *Base case, but IPOs with market value below 1st and above 99th percentile are excluded from the analysis; settings for MV[5-95] and MV[10-90] follow accordingly. Percentiles are based on market value at the end of the first month of trading. In parentheses t-statistics for whether the mean coefficient is equal to zero or one. ***, **, * denotes significance at the one, five, ten percent level or better. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Results of base case are based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24; MV[1-99] observations:1558/1559/1560/1560/1559/1532/1474;MV[5-95]:1432/1433/1434/1434/1434/1409/1359; MV[10-90]: 1272/1273/1274/1274/1274/1253/1210.*

Table 4.15: Robustness Tests of HVZ/EP Model for Different Sample Periods

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
HVZ/EP-rCOMP	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*	Coeff. *(t-stat.)*	Adj. R^2 *(%)*
Base	-0.6254	0.29	-0.1369	0.02	-0.1094	0.07	-0.4316	0.02	0.9403	0.41	1.0383	1.60	0.9573	0.91
= 0	*(-1.51)*		*(-0.35)*		*(-0.29)*		*(-1.01)*		*(1.81)**		*(2.52)***		*(2.69)****	
= 1	*(3.93)****		*(2.88)****		*(2.92)****		*(3.34)****		*(0.10)*		*(0.10)*		*(0.10)*	
1980/89	-0.7203	0.55	-0.4616	0.04	-0.2964	0.21	-0.7275	0.14	1.0208	0.60	1.0981	1.75	0.9789	0.96
= 0	*(-1.29)*		*(-0.79)*		*(-0.36)*		*(-0.90)*		*(0.84)*		*(1.42)*		*(1.84)**	
= 1	*(3.08)****		*(2.51)***		*(1.58)*		*(2.15)***		*(0.00)*		*(0.14)*		*(0.00)*	
1990/99	-0.3501	0.07	0.5373	0.06	0.1517	0.02	-0.2948	0.04	0.8996	0.26	1.7389	2.17	1.1396	0.85
= 0	*(-0.35)*		*(0.57)*		-0.24		*(-0.32)*		*(1.65)*		*(2.84)****		*(1.92)**	
= 1	*(1.36)*		*(0.49)*		*(1.32)*		*(1.40)*		*(0.17)*		*(1.21)*		*(0.25)*	
2000/13	-0.8260	0.57	-0.5253	0.02	-0.1919	0.28	-0.2549	0.11	0.8961	0.12	0.1933	0.15	0.7026	1.94
= 0	*(-1.53)*		*(-1.51)*		(-0.37)		*(-0.63)*		*(0.99)*		*(0.27)*		*(0.85)*	
= 1	*(3.38)****		*(4.40)****		*(2.32)***		*(3.11)****		*(0.10)*		*(1.11)*		*(0.36)*	
Ex.Crises	-0.4850	0.35	-0.2410	0.00	-0.0910	0.02	-0.5608	0.14	0.6945	0.09	0.8298	0.81	0.6909	0.20
Coeff. = 0	*(-1.73)**		*(-0.78)*		(-0.24)		*(-1.40)*		*(1.09)*		*(1.83)**		*(1.84)**	
Coeff. = 1	*(5.29)****		*(4.04)****		*(2.87)****		*(3.89)****		*(0.48)*		*(0.37)*		*(0.82)*	

The table reports FMB regression coefficients of the HVZ/EP model for different sample periods. ***Base:*** *FMB regressions of future realised returns on the composite ICC measure HVZ/EP-rCOMP from 1980 - 2013, where ICC measure set to range between zero and 50 percent, actual future return are used, and at least 15 IPOs per year are required to be included;* ***1980/89:*** *Base case, but IPOs pertaining to fiscal years 1980 to 1989 only; 1990/99 and 2000/13 follows accordingly.* ***Ex.Crises:*** *Base case, but IPOs pertaining to fiscal years 1997/2001 and 2004/07 excluded from sample. In parentheses t-statistics for whether the mean coefficient is equal to zero or one. ***, **, * denotes significance at the one, five, ten percent level or better. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading. Results of base case are based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24; 1980/89 observations: 442/443/444/444/443/439/431; 1990/99: 907/907/907/907/ 907/899/ 863; 2000/13:241/241/241/241/241/225/210; Ex.Crises: 1310/1311/ 1312/1312/1311/1285/1235.*

4.5 Summary and Conclusions

This study elaborates on recent advancements in ICC research and evaluates the performance of mechanical earnings predictions for newly listed firms. IPO markets constitute one of the most demanding empirical settings in finance, and as such are an ideal environment to scrutinise extant "large sample study" results. Using three cross-sectional earnings forecasting models suggested in seminal work by HVZ (2012) and LM (2014), this paper provides comparative evidence for a sample of 1,657 IPOs on the quality of model-based earnings forecasts and the validity of ICC estimates derived therefrom.

My primary two conclusions are as follows. First, combining the HVZ and EP model into one forecasting solution (HVZ/EP) generates higher quality earnings forecasts (as indicated by lower forecast biases and higher ERC coefficients) and more valid ICC estimates (as shown by more consistent FMB coefficients) vis-à-vis the HVZ, EP and RI models stand-alone. These results contrast with large sample evidence of LM (2014)—who recommend that "future research use the RI model or the EP model to generate earnings forecasts" (LM, 2014, p. 1152)—and indicates that model performance is likely to be conditional on firm characteristics: while parsimonious forecasting models (such as EP and RI) appear to work well for established firms, more complex forecasting solutions (such as HVZ/EP) perform better for younger and smaller firms as corroborated by my findings.

Second—and putting findings into greater context—results demonstrate that cross-sectional earnings forecasting models can be reliably used in a market setting where idiosyncratic information is limited and risk factor-based (RFB) estimates are inapplicable due to a lack of sufficiently long enough return histories (i.e., under this study's methodology an IPOs ICC estimate is essentially available from the first day of trading, while RFB proxies calculated from asset pricing models generally require previous five year of return data). On the one hand, this adds evidence to the discussion on the usefulness of IPO prospectus information for investors: pre-IPO financial information is no longer only a valuable source for pricing purposes (Klein, 1996), but can also be effectively used for cost of equity calculations. On the other hand, findings endorse this study's ICC methodology as a unique method of expected return estimation in the context of newly listed firms which, more generally, promotes model-based ICC estimates as a valid alternative to other CoE measures.

While this study contributes novel evidence on the applicability of model-based ICC for newly listed firms, it is silent on the more general impact of firm characteristics and market-settings on the reliability of earnings models and corresponding ICC estimates. Future research might fill this gap and provide clearer guidance on the conditions under which the various models perform best.

4.6 Appendices

Appendix 4.1: Pooled Average Implied Cost of Capital

	MV1		MV3		MV6		MV9		MV12		MV18		MV24	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
rPE														
HVZ	0.0509	0.0337	0.0673	0.0424	0.0677	0.0414	0.0694	0.0411	0.0715	0.0401	0.0881	0.0494	0.0896	0.0464
EP	0.0500	0.0346	0.0668	0.0444	0.0668	0.0430	0.0686	0.0429	0.0715	0.0414	0.0899	0.0496	0.0902	0.0465
RI	0.0509	0.0354	0.0678	0.0447	0.0677	0.0428	0.0695	0.0427	0.0723	0.0411	0.0901	0.0472	0.0902	0.0446
HVZ/EP	0.0444	0.0292	0.0567	0.0360	0.0570	0.0344	0.0584	0.0347	0.0607	0.0329	0.0755	0.0384	0.0762	0.0365
rPEG														
HVZ	0.1255	0.1085	0.1072	0.0918	0.1070	0.0914	0.1081	0.0908	0.1101	0.0921	0.1083	0.0881	0.1080	0.0878
EP	0.1334	0.1185	0.1146	0.0968	0.1142	0.0945	0.1155	0.0951	0.1177	0.0944	0.1246	0.1032	0.1248	0.1017
RI	0.1339	0.1185	0.1145	0.0970	0.1139	0.0945	0.1153	0.0941	0.1175	0.0942	0.1247	0.1047	0.1248	0.1032
HVZ/EP	0.1184	0.1045	0.1079	0.0929	0.1074	0.0924	0.1085	0.0908	0.1105	0.0915	0.1157	0.0960	0.1159	0.0945
rMPEG														
HVZ	0.1334	0.1112	0.1179	0.0951	0.1179	0.0948	0.1189	0.0934	0.1212	0.0956	0.1204	0.0925	0.1200	0.0924
EP	0.1395	0.1208	0.1227	0.1004	0.1225	0.0988	0.1236	0.0977	0.1263	0.0975	0.1345	0.1068	0.1348	0.1065
RI	0.1400	0.1207	0.1223	0.1010	0.1220	0.0984	0.1232	0.0976	0.1259	0.0985	0.1344	0.1081	0.1345	0.1070
HVZ/EP	0.1262	0.1066	0.1179	0.0962	0.1175	0.0955	0.1184	0.0930	0.1209	0.0941	0.1270	0.0992	0.1269	0.0977
rAEGM														
HVZ	0.1430	0.1231	0.1250	0.1027	0.1251	0.1024	0.1260	0.1019	0.1280	0.1034	0.1254	0.0980	0.1247	0.0974
EP	0.1478	0.1301	0.1298	0.1084	0.1297	0.1078	0.1307	0.1066	0.1331	0.1077	0.1395	0.1142	0.1397	0.1124
RI	0.1481	0.1297	0.1300	0.1080	0.1297	0.1081	0.1308	0.1062	0.1333	0.1077	0.1398	0.1148	0.1398	0.1135
HVZ/EP	0.1357	0.1164	0.1253	0.1037	0.1249	0.1029	0.1256	0.1015	0.1280	0.1022	0.1327	0.1066	0.1325	0.1043

The table reports pooled mean and median expected returns the four different implied cost of capital measures suggest by Easton (2004). HVZ, EP, RI and HVZ/EP denote the model which is used to predict earnings. MV1-MV24 denote the market value of equity at the end of the first to the twenty-fourth month of trading which is used to calculate the respective ICC.

Appendix 4.2: Regressions of Future Annual Returns on ICC measures

	MV1	MV3	MV6	MV9	MV12	MV18	MV24
	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*
rPE							
HVZ	-1.0303	-0.1509	0.7744	0.5511	1.5735	0.3774	0.3458
	(-1.61)	*(-0.45)*	*(1.54)*	*(1.25)*	*(2.98)*	*(1.21)*	*(1.39)*
EP	-0.4288	0.2876	0.0218	-0.0866	0.9146	0.7482	0.4908
	(-0.75)	*(0.72)*	*(0.07)*	*(-0.19)*	*(1.67)*	*(2.26)*	*(1.61)*
RI	-0.4030	0.2651	0.0313	-0.0652	0.8721	0.6675	0.3859
	(-0.70)	*(0.68)*	*(0.10)*	*(-0.14)*	*(1.58)*	*(2.29)*	*(1.42)*
HVZ/EP	-0.7889	0.0509	0.0078	-0.2127	0.7712	0.4097	0.3865
	(-1.30)	*(0.13)*	*(0.02)*	*(-0.44)*	*(1.38)*	*(1.31)*	*(1.45)*
rPEG							
HVZ	-0.9230	-0.0334	0.4082	0.2490	1.3266	1.2535	1.0107
	(-2.70)	*(-0.09)*	*(1.00)*	*(0.61)*	*(2.82)*	*(2.51)*	*(2.23)*
EP	-0.6373	0.3734	0.2025	0.1454	1.2124	0.9169	0.7993
	(-1.99)	*(0.90)*	*(0.54)*	*(0.36)*	*(2.59)*	*(1.80)*	*(1.79)*
RI	-0.6558	0.4797	0.1516	0.0087	0.9659	0.9787	0.7878
	(-2.01)	*(1.08)*	*(0.40)*	*(0.02)*	*(2.18)*	*(1.94)*	*(1.70)*
HVZ/EP	-0.9079	0.2010	0.2786	0.0881	1.4829	1.1438	0.9038
	(-2.44)	*(0.51)*	*(0.71)*	*(0.20)*	*(2.42)*	*(2.34)*	*(2.06)*
rMPEG							
HVZ	-0.6596	-0.1476	0.0923	-0.0945	0.7866	1.0910	0.6990
	(-1.93)	*(-0.49)*	*(0.26)*	*(-0.27)*	*(2.00)*	*(2.57)*	*(1.90)*
EP	-0.5755	0.1359	0.0456	-0.0086	1.0096	0.8516	0.6042
	(-1.62)	*(0.35)*	*(0.13)*	*(-0.02)*	*(2.43)*	*(2.00)*	*(1.57)*
RI	-0.6035	0.1713	0.0370	-0.0833	0.9107	0.8740	0.5915
	(-1.68)	*(0.43)*	*(0.10)*	*(-0.24)*	*(2.31)*	*(2.10)*	*(1.50)*
HVZ/EP	-0.6569	-0.0683	-0.0412	-0.1024	0.8957	0.9832	0.6912
	(-1.89)	*(-0.21)*	*(-0.12)*	*(-0.29)*	*(2.25)*	*(2.37)*	*(1.92)*
rAEGM							
HVZ	-0.4862	-0.0182	0.1892	-0.1008	0.8500	1.1178	0.7935
	(-1.42)	*(-0.06)*	*(0.52)*	*(-0.27)*	*(2.02)*	*(2.49)*	*(2.17)*
EP	-0.4931	0.0776	0.1572	0.1391	1.3644	1.2002	0.8154
	(-1.43)	*(0.21)*	*(0.40)*	*(0.35)*	*(2.43)*	*(2.41)*	*(2.16)*
RI	-0.5510	0.0816	0.0773	-0.0137	1.1350	1.1770	0.7925
	(-1.59)	*(0.21)*	*(0.19)*	*(-0.04)*	*(2.33)*	*(2.37)*	*(2.01)*
HVZ/EP	-0.4138	-0.0064	-0.0426	-0.1371	0.9226	1.0353	0.7765
	(-1.14)	*(-0.02)*	*(-0.12)*	*(-0.36)*	*(2.09)*	*(2.37)*	*(2.20)*

The table reports Fama-MacBeth regression coefficients of future realised returns on the four different implied cost of capital measures suggest by Easton (2004). HVZ, EP, RI and HVZ/EP denote the model which is used to predict earnings. MV1 to MV24 denote the market value of equity at the end of the first to twenty-fourth month of trading which is used to calculate the ICC measures. Results based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24.

Appendix 4.3: Future Annual Returns and ICC measures (Deflated Earnings)

Panel A: Pooled average future realised returns and implied cost of capital

	P1		P3		P6		P9		P12		P18		P24	
rCOMP	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
HVZ	0.1113	0.0919	0.1047	0.0834	0.1066	0.0840	0.1091	0.0842	0.1127	0.0865	0.1176	0.0893	0.1202	0.0927
EP	0.1147	0.1003	0.1079	0.0890	0.1098	0.0880	0.1123	0.0885	0.1164	0.0899	0.1308	0.1040	0.1339	0.1048
RI	0.1148	0.0998	0.1071	0.0871	0.1088	0.0869	0.1113	0.0867	0.1154	0.0880	0.1305	0.1040	0.1335	0.1051
HVZ/EP	0.1017	0.0860	0.1006	0.0821	0.1023	0.0814	0.1044	0.0813	0.1079	0.0832	0.1194	0.0933	0.1221	0.0951

Panel B: Regressions of future annual returns on ICC measures

	P1	P3	P6	P9	P12	P18	P24
rCOMP	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*	Coeff. *(t-stat.)*
HVZ	-0.7581	-0.1518	0.3309	0.2035	1.2835	1.0543	0.7791
= 0	*(-1.83)**	*(-0.44)*	*(0.82)*	*(0.54)*	*(2.86)****	*(2.69)***	*(2.25)***
= 1	*(4.24)****	*(3.32)****	*(1.66)*	*(2.10)***	*(0.63)*	*(0.14)*	*(0.64)*
EP	-0.5459	0.0111	-0.0234	-0.2740	1.2017	0.9615	0.6726
= 0	*(-1.21)*	*(0.03)*	*(-0.06)*	*(-0.56)*	*(1.87)**	*(2.30)***	*(1.87)**
= 1	*(3.44)****	*(2.57)***	*(2.65)***	*(2.62)***	*(0.32)*	*(0.10)*	*(0.91)*
RI	-0.5461	0.0143	-0.0096	-0.2531	1.0990	0.9966	0.6819
= 0	*(-1.19)*	*(0.04)*	*(-0.02)*	*(-0.55)*	*(1.79)**	*(2.40)***	*(1.87)**
= 1	*(3.38)****	*(2.52)***	*(2.59)***	*(2.73)***	*(0.17)*	*(0.00)*	*(0.87)*
HVZ/EP	-0.6254	-0.1558	-0.1011	-0.3782	0.9178	0.9968	0.8520
= 0	*(-1.51)*	*(-0.40)*	*(-0.27)*	*(-0.88)*	*(1.78)**	*(2.50)***	*(2.49)***
= 1	*(3.93)****	*(2.95)****	*(2.90)****	*(3.21)****	*(0.17)*	*(0.00)*	*(0.44)*

*Panel A reports pooled mean and median future realised returns and expected returns for the composite ICC measure rCOMP. HVZ, EP, RI and HVZ/EP denote the model which is used to predict earnings. P1 to P24 denote closing stock price at the end of the first to twenty-fourth month of trading which is used to calculate the ICC measures. Expected return observations for all models from P1 to P24:1600/1594/1594/1592/1592/1592/1591. Panel B reports Fama-MacBeth regression coefficients of future realised returns on the composite ICC measure rCOMP. In parentheses t-statistics for whether the mean coefficient is equal to zero or one. ***, **, * denotes significance at the one, five, ten percent level or better. Results based on 1590/1591/1592/1592/1591/1563/1504 observations for MV1 to MV24.*

5 Summary, Conclusion and Future Work

This thesis is motivated by the information-based return models of Easley and O'Hara (2004) and Lambert, Leuz and Verrecchia (2012) who show that firms with higher (lower) quality information environments enjoy relatively lower (higher) cost of equity than otherwise identical firms. More specifically, it is demonstrated that firms can lower their cost of equity if they (1) disclose more value-relevant information to investors (*Quantity*); (2) provide information of higher accuracy (*Precision*); and (3) disseminate information more widely between investor groups (*Asymmetry*). Building on these propositions, I contribute three interrelated papers to this strand of research.

The first paper “Idiosyncratic Information and Expected Rate of Returns: A Meta-Analytic Review of the Literature” provides a quantitative review of the empirical literature examining the association between firm-specific information and expected rate of returns. In total, the results of 113 unique papers—which analyse the cost of equity effects of information *Precision*, *Asymmetry* and *Quantity*—are reconsidered. Overall, findings suggest that the association between firm-specific information and expected returns is moderated by the empirical measurement of both CoE and information attributes: firstly, CoE effects of *Precision*, *Asymmetry* and *Quantity* are 3 to 6 times stronger in studies using risk factor-based (RFB)/valuation model-based (VMB) proxies than in studies conducing asset pricing tests or using realised returns as proxies for CoE; secondly, the controversy over the impact of *Precision* (*Asymmetry*) on firms’ CoE stems—by and large—from the debate on the market pricing of accrual quality (PIN scores) insofar as, when other proxies are used, results confirm the conjectured associations with CoE; similarly, *Quantity* results show that country-level factors and type of information (financial vs. partial-/non-financial) mitigate disclosure effects on firms’ CoE.

However, the variety of different CoE and information proxies used in extant work hinders analysis of the relative importance of *Precision*, *Asymmetry* and *Quantity* as drivers of expected returns; that is, which of the three attributes has comparatively greater CoE relevance cannot be conclusively answered by my literature review. Therefore, in my second paper “The Impact of Idiosyncratic Information on Expected Rate of Returns: A Structural Equation Modelling Approach” I apply a structural equation modelling

(SEM) approach to allow for simultaneous analysis of the different information attributes and CoE measures within one empirical model.

Using three different proxies for each information attribute and nine different CoE measures, SEM results confirm for a sample of 7,091 firms that companies with high (low) quality information environments enjoy relatively lower (higher) CoE than otherwise identical firms, with *Precision* and *Asymmetry* being of equal CoE relevance, while *Quantity* effects being economically negligible; however, findings also show that the significance of this impact decreases with firm size, maturity and profitability as well as market competition. Furthermore, informational differences between companies explain substantial variation in analyst-based implied cost of capital (ICC) estimates, but none in traditional risk factor-based return proxies (RFB), indicating that the former impound much more firm-specific information than the latter. Given the generally higher construct validity of ICC over RFB proxies, this suggests that the incorporation of idiosyncratic information in the measurement of risk factor-based proxies might help improve the empirical soundness of those estimates.

It should be noted, that findings of my literature review as well as SEM paper hinge upon the empirical soundness of VMB measures. While convincing evidence exists that VMB proxies are indeed better measures of CoE than realised return-based ones—which corroborates previous results—, concerns about upward-biased ICC estimates due to optimism in analyst earnings forecasts as well as coverage bias due to a lack of analyst following of young, small and financial distressed firms should not be neglected. My third paper "Implied Cost of Capital and Cross-Sectional Earnings Forecasting Models: Evidence from Newly Listed Firms" addresses these limitations in that it elaborates on recent advancements in model-based ICC research and examines the applicability of this new methodology for the smallest, youngest and least followed firms in capital markets: initial public offerings (IPOs).

Using three cross-sectional earnings forecasting models proposed by Hou et al. (2012, HVZ) and Li and Mohanram (2014, LM), I provide for a sample of 1,657 IPOs comparative evidence on the quality of model-based earnings forecasts and the validity of ICC estimates derived therefrom. Results demonstrate that combining the earnings model of HVZ (2012) with the earnings persistence (EP) model of LM (2014) into one forecasting solution (HVZ/EP) generates less forecast bias, higher earnings response coefficients (ERCs) and more valid ICC estimates vis-à-vis the HVZ, EP and RI (residual income) models stand-alone. This implies that for smaller and younger firms more complex forecasting solutions are required to ensure reliability of model-based earnings predictions and ICC estimates.

Findings also indicate some interesting avenues for future research. First, quantifying the effect of idiosyncratic information on firms' CoE via an information risk factor, which contemporaneously captures firms' asymmetry and precision characteristics, appears to be promising in that it might help improve the empirical soundness of asset pricing models and, as such, expands recent work by Ecker et al. (2006), who recommend to modify traditional RFB models by an earnings quality factor. Second, investigating the impact of firm characteristics and market settings on both the reliability of mechanical earnings forecasts and the validity of model-based ICC estimates might be beneficial to obtain clearer guidance on the conditions under which the respective models perform best; that is, while parsimonious forecasting models appear to work well for established firms and more complex forecasting solutions tend to perform better for younger and smaller firms; the drivers of model performance are unidentified in extant work.

The practical implications of this thesis may be summarised as follows: (1) more corporate disclosure is no guarantee for lower cost of equity *per se*; that is, disclosure effects are greatest if financial information is disclosed, but starkly diminished for partial/non-financial disclosure; furthermore, firms' operating in countries with weak disclosure regulation and transparency enjoy substantially greater CoE benefits than firms operating in highly regulated markets; therefore, firms should consider these points before committing to more expansive disclosure policies given the economic trade-off between the proprietary costs of more disclosure and the benefits from lower costs of equity; (2) if firms must decide to either provide more accurate information to investors or distribute information more equally between them, they should dedicate scarce resources to those activities that show the most room for improvement given similarly strong CoE effects emanating from

Asymmetry and *Precision* activities (such activities may include promoting investor relations, increasing reporting quality/frequency, committing to more expansive disclosure); however, if less established firms face a trade-off between providing better quality information to investors and aiming to reduce informational disadvantages between investor groups, they should choose the former over the latter given the greater persistence of *Precision* effects (i.e., CoE benefits) for younger and smaller firms; (3) firms may operationalise model-based ICC estimates as a valid alternative to commonly used return proxies; while particularly relevant for firms for which traditional CoE measures are unavailable (due to a lack of a trading history or no listing at all), this approach also offers public companies a simple method of cross-validating their existing cost of equity estimates.

Overall, this thesis makes three major contributions to the literature. First, providing a quantitative review of the literature on the association between idiosyncratic information and expected returns complements narrative reviews on this topic. Second, reassessing this association by means of a novel structural equation modelling approach contributes new and empirically robust evidence to the debate on the market pricing of information risk. Third, examining the applicability of cross-sectional earnings forecasting models in an imperfect market setting is unique and elaborates on recent advancements in model-based ICC research.

Taken all findings together, this thesis presents convincing evidence that firms can lower their cost of equity if they disclose more value-relevant information to investors, provide information of higher accuracy, and disseminate information more widely between investor groups. While there is debate about the economic and statistical significance of these associations, results suggest that it is limited to the controversy over the market pricing of a few information proxies. Reconsidering a large sample of previous results and using a wide range of different measures for both CoE and information attributes, findings corroborate the proposition that firms with high (low) quality information environments enjoy relatively low (high) costs of equity.

List of References

ABOODY, D., HUGHES, J. & LIU, J. 2005. Earnings Quality, Insider Trading, and Cost of Capital. *Journal of Accounting Research,* 43**,** 651-673.

ADMATI, A. R. 1985. A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets. *Econometrica,* 53**,** 629-657.

AHMED, K. & COURTIS, J. K. 1999. Associations between corporate characteristics and disclosure levels in annual reports: a meta-analysis. *The British Accounting Review,* 31**,** 35-61.

AIMR. 2001. *Fair but not full disclosure: Everyone has less information under regulation FD, analysts, portfolio managers say in AIMR survey* [Online]. Available: https://www.cfainstitute.org/about/press/release/Documents/2001_press_releases.pdf [Accessed 01/04/2016].

AKINS, B. K., JEFFREY, N. & VERDI, R. S. 2012. Investor Competition over Information and the Pricing of Information Asymmetry. *The Accounting Review,* 87**,** 35-58.

AL-HADI, A., TAYLOR, G. & HOSSAIN, M. 2015. Disaggregation, auditor conservatism and implied cost of equity capital: An international evidence from the GCC. *Journal of Multinational Financial Management,* 29**,** 66-98.

AL GUINDY, M. 2016. Corporate Twitter Use and Cost of Capital. Available at SSRN: https://ssrn.com/abstract=2824733.

AMIHUD, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets,* 5**,** 31-56.

AMIHUD, Y. & MENDELSON, H. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics,* 17**,** 223-249.

ANG, A., HODRICK, R. J., XING, Y. & ZHANG, X. 2006. The Cross-Section of Volatility and Expected Returns. *The Journal of Finance,* 61**,** 259-299.

ARMSTRONG, C. S., CORE, J. E., TAYLOR, D. J. & VERRECCHIA, R. E. 2011. When Does Information Asymmetry Affect the Cost of Capital? *Journal of Accounting Research,* 49**,** 1-40.

ARTIACH, T. C. & CLARKSON, P. M. 2011. Disclosure, conservatism and the cost of equity capital: A review of the foundation literature. *Accounting and Finance,* 51**,** 2-49.

ARTIACH, T. C. & CLARKSON, P. M. 2014. Conservatism, disclosure and the cost of equity capital. *Australian Journal of Management,* 39**,** 293-314.

ASLAN, H., EASLEY, D., HVIDKJAER, S. & O'HARA, M. 2011. The characteristics of informed trading: Implications for asset pricing. *Journal of Empirical Finance,* 18**,** 782-801.

BACHOO, K., TAN, R. & WILSON, M. 2013. Firm Value and the Quality of Sustainability Reporting in Australia. *Australian Accounting Review,* 23**,** 67-87.

BAGEHOT, W. 1971. The Only Game in Town. *Financial Analysts Journal,* 27**,** 12-22.

BAGINSKI, S. P. & RAKOW, K. C. 2012. Management earnings forecast disclosure policy and the cost of equity capital. *Review of Accounting Studies,* 17**,** 279-321.

BALL, R. & BROWN, P. 1968. An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research,* 6**,** 159-178.

BARINOV, A. 2013. Analyst disagreement and aggregate volatility risk. *Journal of Financial and Quantitative Analysis,* 48**,** 1877-1900.

BARRON, O. E., BYARD, D. & KIM, O. 2002. Changes in Analysts' Information around Earnings Announcements. *The Accounting Review,* 77**,** 821-846.

BARRON, O. E., HARRIS, D. G. & MARY, S. 2005. Evidence That Investors Trade on Private Event-Period Information around Earnings Announcements. *The Accounting Review,* 80**,** 403-421.

BARRON, O. E., KIM, O., LIM, S. C. & STEVENS, D. E. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review***,** 421-433.

BARRON, O. E., SHENG, X. S. & THEVENOT, M. 2012. The Information Environment and Cost of Capital. Available at SSRN: https://ssrn.com/abstract=2099825.

BARRON, O. E. & STUERKE, P. S. 1998. Dispersion in Analysts' Earnings Forecasts as a Measure of Uncertainty. *Journal of Accounting, Auditing & Finance,* 13**,** 245-270.

BARRY, C. B. & BROWN, S. J. 1984. Differential information and the small firm effect. *Journal of Financial Economics,* 13**,** 283-294.

BARRY, C. B. & BROWN, S. J. 1985. Differential Information and Security Market Equilibrium. *Journal of Financial and Quantitative Analysis,* 20**,** 407-422.

BARTH, M. E., BEAVER, W. H. & LANDSMAN, W. R. 2001. The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of Accounting and Economics,* 31**,** 77-104.

BARTH, M. E. & CLINCH, G. 2009. Scale Effects in Capital Markets-Based Accounting Research. *Journal of Business Finance & Accounting,* 36**,** 253-288.

BARTH, M. E., KONCHITCHKI, Y. & LANDSMAN, W. R. 2013. Cost of capital and earnings transparency. *Journal of Accounting and Economics,* 55**,** 206-224.

BARTOV, E., MOHANRAM, P. & SEETHAMRAJU, C. 2002. Valuation of Internet Stocks—An IPO Perspective. *Journal of Accounting Research,* 40**,** 321-346.

BEATTIE, V., MCINNES, B. & FEARNLEY, S. 2002. Through the eyes of management: a study of narrative disclosures. London: Centre for Business Performance, The Institute of Chartered Accountants in England and Wales.

BEAVER, W. H. 1968. The Information Content of Annual Earnings Announcements. *Journal of Accounting Research,* 6**,** 67-92.

BENEISH, M. D., BILLINGS, M. B. & HODDER, L. D. 2008. Internal control weaknesses and information uncertainty. *The Accounting Review,* 83**,** 665-703.

BERETTA, S. & BOZZOLAN, S. 2004. A framework for the analysis of firm risk communication. *The International Journal of Accounting,* 39**,** 265-288.

BERETTA, S. & BOZZOLAN, S. 2008. Quality versus Quantity: The Case of Forward-Looking Disclosure. *Journal of Accounting, Auditing & Finance,* 23**,** 333-376.

BERGER, P. G., CHEN, H. J. & LI, F. 2012. Firm specific information and the cost of equity capital. Available at SSRN: https://ssrn.com/abstract=906152.

BERTOMEU, J., BEYER, A. & DYE, R. A. 2011. Capital Structure, Cost of Capital, and Voluntary Disclosures. *The Accounting Review,* 86**,** 857-886.

BEYER, A., COHEN, D. A., LYS, T. Z. & WALTHER, B. R. 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics,* 50**,** 296-343.

BHATTACHARYA, N., DESAI, H. & VENKATARAMAN, K. 2013. Does Earnings Quality Affect Information Asymmetry? Evidence from Trading Costs*. *Contemporary Accounting Research,* 30**,** 482-516.

BHATTACHARYA, N., ECKER, F., OLSSON, P. M. & SCHIPPER, K. 2012. Direct and Mediated Associations among Earnings Quality, Information Asymmetry, and the Cost of Equity. *The Accounting Review,* 87**,** 449-482.

BHATTACHARYA, U., DAOUK, H. & WELKER, M. 2003. The world price of earnings opacity. *The Accounting Review,* 78**,** 641-678.

BIDDLE, G. C., SEOW, G. S. & SIEGEL, A. F. 1995. Relative versus Incremental Information Content. *Contemporary Accounting Research,* 12**,** 1-23.

BLANCO, B., GARCIA LARA, J. M. & TRIBO, J. A. 2015. Segment Disclosure and Cost of Capital. *Journal of Business Finance & Accounting,* 42**,** 367-411.

BOEHMER, E., GRAMMIG, J. & THEISSEN, E. 2007. Estimating the probability of informed trading—does trade misclassification matter? *Journal of Financial Markets,* 10**,** 26-47.

BOTOSAN, C. A. 1997. Disclosure Level and the Cost of Equity Capital. *The Accounting Review,* 72**,** 323-349.

BOTOSAN, C. A. & PLUMLEE, M. A. 2002. A Re-examination of Disclosure Level and the Expected Cost of Equity Capital. *Journal of Accounting Research,* 40**,** 21-40.

BOTOSAN, C. A. & PLUMLEE, M. A. 2005. Assessing Alternative Proxies for the Expected Risk Premium. *The Accounting Review,* 80**,** 21-53.

BOTOSAN, C. A. & PLUMLEE, M. A. 2013. Are Information Attributes Priced? *Journal of Business Finance & Accounting,* 40**,** 1045-1067.

BOTOSAN, C. A., PLUMLEE, M. A. & WEN, H. E. 2011. The Relation between Expected Returns, Realized Returns, and Firm Risk Characteristics. *Contemporary Accounting Research,* 28**,** 1085-1122.

BOTOSAN, C. A., PLUMLEE, M. A. & XIE, Y. 2004. The Role of Information Precision in Determining the Cost of Equity Capital. *Review of Accounting Studies,* 9**,** 233-259.

BOUJELBENE, M. A. & AFFES, H. 2013. The impact of intellectual capital disclosure on cost of equity capital: A case of French firms. *Journal of Economics Finance and Administrative Science,* 18**,** 45-53.

BRADSHAW, M., DRAKE, M., MYERS, J. & MYERS, L. 2012. A re-examination of analysts' superiority over time-series forecasts of annual earnings. *Review of Accounting Studies,* 17**,** 944-968.

BRAV, A. & GOMPERS, P. A. 2000. Insider trading subsequent to initial public offerings: Evidence from expirations of lock-up provisions. *Available at SSRN 204094*.

BRAV, A. & GOMPERS, P. A. 2003. The Role of Lockups in Initial Public Offerings. *Review of Financial Studies,* 16**,** 1-29.

BRENNAN, M. J., CHORDIA, T. & SUBRAHMANYAM, A. 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns1. *Journal of Financial Economics,* 49**,** 345-373.

BRENNAN, M. J., HUH, S. W. & SUBRAHMANYAM, A. 2016. Asymmetric Effects of Informed Trading on the Cost of Equity Capital. *Management Science,* 62**,** 2460-2480.

BROWN, S. & HILLEGEIST, S. 2007. How disclosure quality affects the level of information asymmetry. *Review of Accounting Studies,* 12**,** 443-477.

BROWN, S., HILLEGEIST, S. A. & LO, K. 2004. Conference calls and information asymmetry. *Journal of Accounting and Economics,* 37**,** 343-366.

BUSHMAN, R., CHEN, Q., ENGEL, E. & SMITH, A. 2004. Financial accounting information, organizational complexity and corporate governance systems. *Journal of Accounting and Economics,* 37**,** 167-201.

BYARD, D. & SHAW, K. W. 2003. Corporate Disclosure Quality and Properties of Analysts' Information Environment. *Journal of Accounting, Auditing & Finance,* 18**,** 355-378.

CALLEN, J. L., KHAN, M. & LU, H. 2013. Accounting Quality, Stock Price Delay, and Future Stock Returns. *Contemporary Accounting Research,* 30**,** 269-295.

CAMPBELL, J. & SHILLER, R. 1988a. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies,* 1**,** 195-228.

CAMPBELL, J. L., CHEN, H., DHALIWAL, D. S., LU, H.-M. & STEELE, L. B. 2014. The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies,* 19**,** 396-455.

CAMPBELL, J. Y. & SHILLER, R. J. 1988b. Stock prices, earnings, and expected dividends. *The Journal of Finance,* 43**,** 661-676.

CAO, Y., MYERS, J. N., MYERS, L. A. & OMER, T. C. 2015. Company reputation and the cost of equity capital. *Review of Accounting Studies,* 20**,** 42-81.

CAO, Y., MYERS, L. A., TSANG, A. & YANG, Y. G. 2017. Management forecasts and the cost of equity capital: international evidence. *Review of Accounting Studies***,** 1-48.

CARHART, M. M. 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance,* 52**,** 57-82.

CHAN, A. L. C., LIN, S. W. J. & STRONG, N. 2009. Accounting conservatism and the cost of equity capital: UK evidence. *Managerial Finance,* 35**,** 325-345.

CHAN, C. Y., LO, H. C. & YANG, M. J. 2016. The revision frequency of earnings forecasts and firm characteristics. *North American Journal of Economics and Finance,* 35**,** 116-132.

CHARITOU, A., CLUBB, C. & ANDREOU, A. 2001. The Effect of Earnings Permanence, Growth, and Firm Size on the Usefulness of Cash Flows and Earnings in Explaining Security Returns: Empirical Evidence for the UK. *Journal of Business Finance & Accounting,* 28**,** 563-594.

CHEN, K. C., WEI, K.-C. J. & CHEN, Z. 2003. Disclosure, corporate governance, and the cost of equity capital: evidence from Asia's emerging markets. Available at SSRN: https://ssrn.com/abstract=422000.

CHEN, L., DA, Z. & ZHAO, X. 2013. What Drives Stock Price Movements? *Review of Financial Studies,* 26**,** 841-876.

CHEN, L. H., DHALIWAL, D. S. & TROMBLEY, M. A. 2008. The Effect of Fundamental Risk on the Market Pricing of Accruals Quality. *Journal of Accounting, Auditing & Finance,* 23**,** 471-492.

CHEN, S., SHEVLIN, T. & TONG, Y. H. 2007. Does the Pricing of Financial Reporting Quality Change Around Dividend Changes? *Journal of Accounting Research,* 45**,** 1-40.

CHEN, Z. H., DHALIWAL, D. S. & XIE, H. 2010. Regulation fair disclosure and the cost of equity capital. *Review of Accounting Studies,* 15**,** 106-144.

CHENG, C. S. A., COLLINS, D. & HUANG, H. H. 2006. Shareholder rights, financial disclosure and the cost of equity capital. *Review of Quantitative Finance and Accounting,* 27**,** 175-204.

CHEYNEL, E. 2013. A theory of voluntary disclosure and cost of capital. *Review of Accounting Studies,* 18**,** 987-1020.

CHIEN, Y. T. & LU, H. M. 2015. Firm websites and the risk of firm. *Industrial Management & Data Systems,* 115**,** 504-520.

CHOI, J. J., JIN, L. & YAN, H. 2016. Informed trading and expected returns. Available at SSRN: https://ssrn.com/abstract=2193733.

CHRISTENSEN, P. O., DE LA ROSA, L. E. & FELTHAM, G. A. 2010. Information and the Cost of Capital: An Ex Ante Perspective. *The Accounting Review,* 85**,** 817-848.

CLARKSON, P., GUEDES, J. & THOMPSON, R. 1996. On the diversification, observability, and measurement of estimation risk. *Journal of Financial and Quantitative Analysis,* 31**,** 69-84.

CLARKSON, P. M., FANG, X., LI, Y. & RICHARDSON, G. 2013. The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy,* 32**,** 410-431.

CLARKSON, P. M. & SATTERLY, A. 1997. Australian evidence on the pricing of estimation risk. *Pacific-Basin Finance Journal,* 5**,** 281-299.

CLARKSON, P. M. & THOMPSON, R. E. X. 1990. Empirical Estimates of Beta When Investors Face Estimation Risk. *The Journal of Finance,* 45**,** 431-453.

CLAUS, J. & THOMAS, J. 2001. Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets. *The Journal of Finance,* 56**,** 1629-1666.

CLEMENT, M., FRANKEL, R. & MILLER, J. 2003. Confirming Management Earnings Forecasts, Earnings Uncertainty, and Stock Returns. *Journal of Accounting Research,* 41**,** 653-679.

COHEN, D. A. 2003. Quality of financial reporting choice: determinants and economic consequences. Available at SSRN: https://ssrn.com/abstract=422581.

COHEN, D. A. 2008. Does information risk really matter? An analysis of the determinants and economic consequences of financial reporting quality. *Asia-Pacific Journal of Accounting & Economics,* 15**,** 69-90.

COLLER, M. & YOHN, T. L. 1997. Management Forecasts and Information Asymmetry: An Examination of Bid-Ask Spreads. *Journal of Accounting Research,* 35**,** 181-191.

COPELAND, T. E. & GALAI, D. 1983. Information Effects on the Bid-Ask Spread. *The Journal of Finance,* 38**,** 1457-1469.

CORE, J. E., GUAY, W. R. & VERDI, R. 2008. Is accruals quality a priced risk factor? *Journal of Accounting and Economics,* 46**,** 2-22.

CORE, J. E., HAIL, L. & VERDI, R. S. 2015. Mandatory disclosure quality, inside ownership, and cost of capital. *European Accounting Review,* 24**,** 1-29.

CORNELLI, F. & GOLDREICH, D. 2003. Bookbuilding: How Informative Is the Order Book? *The Journal of Finance,* 58**,** 1415-1443.

CORNELLI, F., GOLDREICH, D. & LJUNGQVIST, A. 2006. Investor Sentiment and Pre-IPO Markets. *The Journal of Finance,* 61**,** 1187-1216.

CORNELLI, F. & YILMAZ, B. 2015. Do short-selling constraints matter. Available at: http://economics.mit.edu/files/10543.

CORWIN, S. A. & SCHULTZ, P. 2012. A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices. *The Journal of Finance,* 67**,** 719-760.

COX, C. T. 1985. Further evidence on the representativeness of management earnings forecasts. *Accounting Review***,** 692-701.

DECHOW, P., GE, W. & SCHRAND, C. 2010. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics,* 50**,** 344-401.

DECHOW, P. M. & DICHEV, I. D. 2002. The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors. *The Accounting Review,* 77**,** 35-59.

DÉJEAN, F. & MARTINEZ, I. 2009. Environmental Disclosure and the Cost of Equity: The French Case. *Accounting in Europe,* 6**,** 57-80.

DHALIWAL, D., LI, O. Z., TSANG, A. & YANG, Y. G. 2014. Corporate social responsibility disclosure and the cost of equity capital: The roles of stakeholder orientation and financial transparency. *Journal of Accounting and Public Policy,* 33**,** 328-355.

DHALIWAL, D. S., LI, O. Z., TSANG, A. & YANG, Y. G. 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review,* 86**,** 59-100.

DIAMOND, D. W. & VERRECCHIA, R. E. 1991. Disclosure, Liquidity, and the Cost of Capital. *The Journal of Finance,* 46**,** 1325-1359.

DIETHER, K. B., MALLOY, C. J. & SCHERBINA, A. 2002. Differences of Opinion and the Cross Section of Stock Returns. *The Journal of Finance,* 57**,** 2113-2141.

DOUKAS, J. A., KIM, C. F. & PANTZALIS, C. 2006. Divergence of opinion and equity returns. *Journal of Financial and Quantitative Analysis,* 41**,** 573-606.

DUARTE, J., HAN, X., HARFORD, J. & YOUNG, L. 2008. Information asymmetry, information dissemination and the effect of regulation FD on the cost of capital. *Journal of Financial Economics,* 87**,** 24-44.

DUARTE, J. & YOUNG, L. 2009. Why is PIN priced? *Journal of Financial Economics,* 91**,** 119-138.

DUGAR, A. & NATHAN, S. 1995. The Effect of Investment Banking Relationships on Financial Analysts' Earnings Forecasts and Investment Recommendations. *Contemporary Accounting Research,* 12**,** 131-160.

DUTTA, S. & NEZLOBIN, A. 2017. Information disclosure, firm growth, and the cost of capital. *Journal of Financial Economics*, 123**,** 415-431.

EASLEY, D., HVIDKJAER, S. & O'HARA, M. 2010. Factoring Information into Returns. *The Journal of Financial and Quantitative Analysis,* 45**,** 293-309.

EASLEY, D., HVIDKJAER, S. & O’HARA, M. 2002. Is Information Risk a Determinant of Asset Returns? *The Journal of Finance,* 57**,** 2185-2221.

EASLEY, D., KIEFER, N. M. & O'HARA, M. 1997. One day in the life of a very common stock. *Review of Financial Studies,* 10**,** 805-835.

EASLEY, D., KIEFER, N. M., O'HARA, M. & PAPERMAN, J. B. 1996. Liquidity, Information, and Infrequently Traded Stocks. *The Journal of Finance,* 51**,** 1405-1436.

EASLEY, D. & O'HARA, M. 2004. Information and the Cost of Capital. *The Journal of Finance,* 59**,** 1553-1583.

EASTON, P. D. 2004. PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital. *The Accounting Review,* 79**,** 73-95.

EASTON, P. D. & MONAHAN, S. J. 2005. An Evaluation of Accounting-Based Measures of Expected Returns. *The Accounting Review,* 80**,** 501-538.

EASTON, P. D. & MONAHAN, S. J. 2016. Review of Recent Research on Improving Earnings Forecasts and Evaluating Accounting-based Estimates of the Expected Rate of Return on Equity Capital. *Abacus,* 52**,** 35-58.

EASTON, P. D. & SOMMERS, G. A. 2003. Scale and the Scale Effect in Market-based Accounting Research. *Journal of Business Finance & Accounting,* 30**,** 25-56.

ECHTERLING, F., EIERLE, B. & KETTERER, S. 2015. A review of the literature on methods of computing the implied cost of capital. *International Review of Financial Analysis,* 42**,** 235-252.

ECKER, F. 2014. Information Precision and Long-Run Performance of Initial Public Offerings. *Contemporary Accounting Research,* 31**,** 876-910.

ECKER, F., FRANCIS, J., KIM, I., OLSSON, P. M. & SCHIPPER, K. 2006. A Returns-Based Representation of Earnings Quality. *The Accounting Review,* 81**,** 749-780.

EL GHOUL, S., GUEDHAMI, O., NI, Y., PITTMAN, J. & SAADI, S. 2013. Does Information Asymmetry Matter to Equity Pricing? Evidence from Firms' Geographic Location*. *Contemporary Accounting Research,* 30**,** 140-181.

ELESWARAPU, V. R. & REINGANUM, M. R. 1993. The seasonal behavior of the liquidity premium in asset pricing. *Journal of Financial Economics,* 34**,** 373-386.

ELIWA, Y., HASLAM, J. & ABRAHAM, S. 2016. The association between earnings quality and the cost of equity capital: Evidence from the UK. *International Review of Financial Analysis,* 48**,** 125-139.

ELTON, E. J. 1999. Presidential Address: Expected Return, Realized Return, and Asset Pricing Tests. *The Journal of Finance,* 54**,** 1199-1220.

ELZAHAR, H., HUSSAINEY, K., MAZZI, F. & TSALAVOUTAS, I. 2015. Economic consequences of key performance indicators' disclosure quality. *International Review of Financial Analysis,* 39**,** 96-112.

EMBONG, Z., MOHD-SALEH, N. & HASSAN, M. S. 2012. Firm size, disclosure and cost of equity capital. *Asian Review of Accounting,* 20**,** 119-139.

ESPINOSA, M. & TROMBETTA, M. 2007. Disclosure interactions and the cost of equity capital: Evidence from the Spanish continuous market. *Journal of Business Finance & Accounting,* 34**,** 1371-1392.

EUGSTER, F. 2014. Endogeneity and the Dynamics of Voluntary Disclosure Quality: Is There Really an Effect on the Cost of Equity Capital? Available at SSRN: https://ssrn.com/abstract=2379683.

EVANS, M. E. 2016. Commitment and Cost of Equity Capital: An Examination of Timely Balance Sheet Disclosure in Earnings Announcements. *Contemporary Accounting Research,* 33**,** 1136-1171.

FAMA, E. F. 1965. The Behavior of Stock-Market Prices. *The Journal of Business,* 38**,** 34-105.

FAMA, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance,* 25**,** 383-417.

FAMA, E. F. & FRENCH, K. R. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics,* 33**,** 3-56.

FAMA, E. F. & FRENCH, K. R. 1997. Industry costs of equity. *Journal of Financial Economics,* 43**,** 153-193.

FAMA, E. F. & FRENCH, K. R. 2000. Forecasting Profitability and Earnings. *The Journal of Business,* 73**,** 161-175.

FAMA, E. F. & FRENCH, K. R. 2004. New lists: Fundamentals and survival rates. *Journal of Financial Economics,* 73**,** 229-269.

FAMA, E. F. & FRENCH, K. R. 2006. Profitability, investment and average returns. *Journal of Financial Economics,* 82**,** 491-518.

FAMA, E. F. & FRENCH, K. R. 2015. A five-factor asset pricing model. *Journal of Financial Economics,* 116**,** 1-22.

FAN, X. & SIVO, S. A. 2005. Sensitivity of Fit Indexes to Misspecified Structural or Measurement Model Components: Rationale of Two-Index Strategy Revisited. *Structural Equation Modeling: A Multidisciplinary Journal,* 12**,** 343-367.

FAN, X., THOMPSON, B. & WANG, L. 1999. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal,* 6**,** 56-83.

FELTHAM, G. A. & OHLSON, J. A. 1995. Valuation and Clean Surplus Accounting for Operating and Financial Activities*. *Contemporary Accounting Research,* 11**,** 689-731.

FELTHAM, G. A. & OHLSON, J. A. 1996. Uncertainty resolution and the theory of depreciation measurement. *Journal of Accounting Research***,** 209-234.

FENG, M. 2014. Discussion of “Evaluating cross-sectional forecasting models for implied cost of capital”. *Review of Accounting Studies,* 19**,** 1186-1190.

FENG, M., LI, C. & GU, Z. 2009. Management Forecasts and the Cost of Equity. Available at SSRN: https://ssrn.com/abstract=1471058.

FIELD, L. C. & HANKA, G. 2001. The Expiration of IPO Share Lockups. *The Journal of Finance,* 56**,** 471-500.

FOSTER, F. D. & VISWANATHAN, S. 1994. Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information. *Journal of Financial and Quantitative Analysis,* 29**,** 499-518.

FRANCIS, J., LAFOND, R., OLSSON, P. & SCHIPPER, K. 2005a. The market pricing of accruals quality. *Journal of Accounting and Economics,* 39**,** 295-327.

FRANCIS, J., LAFOND, R., OLSSON, P. M. & SCHIPPER, K. 2004. Costs of Equity and Earnings Attributes. *The Accounting Review,* 79**,** 967-1010.

FRANCIS, J., NANDA, D. & OLSSON, P. E. R. 2008. Voluntary Disclosure, Earnings Quality, and Cost of Capital. *Journal of Accounting Research,* 46**,** 53-99.

FRANCIS, J. & PHILBRICK, D. 1993. Analysts' Decisions As Products of a Multi-Task Environment. *Journal of Accounting Research,* 31**,** 216-230.

FRANCIS, J., SCHIPPER, K. & VINCENT, L. 2003. The relative and incremental explanatory power of earnings and alternative (to earnings) performance measures for returns. *Contemporary Accounting Research,* 20**,** 121.

FRANCIS, J. R., KHURANA, I. K. & PEREIRA, R. 2005b. Disclosure incentives and effects on cost of capital around the world. *The Accounting Review,* 80**,** 1125-1162.

FU, R., KRAFT, A. & ZHANG, H. 2012. Financial reporting frequency, information asymmetry, and the cost of equity. *Journal of Accounting and Economics,* 54**,** 132-149.

GARCÍA-SÁNCHEZ, I.-M. & NOGUERA-GÁMEZ, L. 2017. Integrated information and the cost of capital. *International Business Review,* 26**,** 959-975.

GARCIA LARA, J. M., GARCIA OSMA, B. & PENALVA, F. 2011. Conditional conservatism and cost of capital. *Review of Accounting Studies,* 16**,** 247-271.

GEBHARDT, W. R., LEE, C. M. C. & SWAMINATHAN, B. 2001. Toward an Implied Cost of Capital. *Journal of Accounting Research,* 39**,** 135-176.

GIETZMANN, M. & IRELAND, J. 2005. Cost of capital, strategic disclosures and accounting choice. *Journal of Business Finance & Accounting,* 32**,** 599-634.

GLOSTEN, L. R. & HARRIS, L. E. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics,* 21**,** 123-142.

GLOSTEN, L. R. & MILGROM, P. R. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics,* 14**,** 71-100.

GODE, D. & MOHANRAM, P. 2003. Inferring the Cost of Capital Using the Ohlson–Juettner Model. *Review of Accounting Studies,* 8**,** 399-431.

GOMES, A., GORTON, G. & MADUREIRA, L. 2007. SEC Regulation Fair Disclosure, information, and the cost of capital. *Journal of Corporate Finance,* 13**,** 300-334.

GORDON, J. R. & GORDON, M., J. 1997. The Finite Horizon Expected Return Model. *Financial Analysts Journal,* 53**,** 52-61.

GRAY, P., KOH, P.-S. & TONG, Y. H. 2009. Accruals Quality, Information Risk and Cost of Capital: Evidence from Australia. *Journal of Business Finance & Accounting,* 36**,** 51-72.

GRÜNING, M. 2011. Capital Market Implications of Corporate Disclosure: German Evidence. *Business Research,* 4**,** 48-72.

GUAY, W., KOTHARI, S. & SHU, S. 2011. Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management,* 36**,** 125-149.

HAIL, L. 2002. The impact of voluntary corporate disclosures on the ex-ante cost of capital for Swiss firms. *European Accounting Review,* 11**,** 741-773.

HANLEY, K. W. 1993. The underpricing of initial public offerings and the partial adjustment phenomenon. *Journal of Financial Economics,* 34**,** 231-250.

HAYN, C. 1995. The information content of losses. *Journal of Accounting and Economics,* 20**,** 125-153.

HE, W. P., LEPONE, A. & LEUNG, H. 2013. Information asymmetry and the cost of equity capital. *International Review of Economics & Finance,* 27**,** 611-620.

HEALY, P. M., HUTTON, A. P. & PALEPU, K. G. 1999. Stock Performance and Intermediation Changes Surrounding Sustained Increases in Disclosure. *Contemporary Accounting Research,* 16**,** 485-520.

HEALY, P. M. & PALEPU, K. G. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics,* 31**,** 405-440.

HEFLIN, F. L., SHAW, K. W. & WILD, J. J. 2005. Disclosure Policy and Market Liquidity: Impact of Depth Quotes and Order Sizes*. *Contemporary Accounting Research,* 22**,** 829-865.

HINSON, L. A. & UTKE, S. 2016. Structural Equation Modeling in Archival Accounting Research: An Application to Disclosure and Cost of Capital.

HIRSHLEIFER, D., HOU, K. & TEOH, S. H. 2012. The accrual anomaly: risk or mispricing? *Management Science,* 58**,** 320-335.

HOLTHAUSEN, R. W. & WATTS, R. L. 2001. The relevance of the value-relevance literature for financial accounting standard setting. *Journal of Accounting and Economics,* 31**,** 3-75.

HONG, H., LIM, T. & STEIN, J. C. 2000. Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance,* 55**,** 265-295.

HOU, K. & ROBINSON, D. T. 2006. Industry concentration and average stock returns. *The Journal of Finance,* 61**,** 1927-1956.

HOU, K., VAN DIJK, M. A. & ZHANG, Y. 2012. The implied cost of capital: A new approach. *Journal of Accounting and Economics,* 53**,** 504-526.

HOU, Y. 2015. The role of diversification in the pricing of accruals quality. *Review of Accounting Studies,* 20**,** 1059-1092.

HRIBAR, P. & COLLINS, D. W. 2002. Errors in Estimating Accruals: Implications for Empirical Research. *Journal of Accounting Research,* 40**,** 105-134.

HU, L. T. & BENTLER, P. M. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal,* 6**,** 1-55.

HUANG, R. D. & STOLL, H. R. 1996. Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics,* 41**,** 313-357.

HUGHES, J. S., LIU, J. & LIU, J. 2007. Information asymmetry, diversification, and cost of capital. *The Accounting Review,* 82**,** 705-729.

HUNTER, J. E., SCHMIDT, F. L. & JACKSON, G. B. 1982. *Meta-Analysis: Cumulating Research Findings Across Studies,* Berverly Hills, SAGE Publications.

HWANG, L.-S., LEE, W.-J., LIM, S.-Y. & PARK, K.-H. 2013. Does information risk affect the implied cost of equity capital? An analysis of PIN and adjusted PIN. *Journal of Accounting & Economics,* 55**,** 148-167.

HWANG, L.-S. & LIM, S.-Y. 2012. Do Investors Price Accruals Quality? A Reexamination in the Implied Cost of Equity Capital. *Asia-Pacific Journal of Financial Studies,* 41**,** 458-490.

IMHOFF, E. A. 1978. The representativeness of management earnings forecasts. *The Accounting Review***,** 836-850.

JEGADEESH, N. & TITMAN, S. 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance,* 48**,** 65-91.

KANG, M. 2010. Probability of information-based trading and the January effect. *Journal of Banking & Finance,* 34**,** 2985-2994.

KHAN, M. 2008. Are accruals mispriced? Evidence from tests of an Intertemporal Capital Asset Pricing Model. *Journal of Accounting and Economics,* 45**,** 55-77.

KHLIF, H. & CHALMERS, K. 2015. A review of meta-analytic research in accounting. *Journal of Accounting Literature,* 35**,** 1-27.

KHLIF, H., SAMAHA, K. & AZZAM, I. 2015. Disclosure, ownership structure, earnings announcement lag and cost of equity capital in emerging markets: The case of the Egyptian stock exchange. *Journal of Applied Accounting Research,* 16**,** 28-57.

KIM, D. & QI, Y. 2010. Accruals quality, stock returns, and macroeconomic conditions. *The Accounting Review,* 85**,** 937-978.

KIM, J.-B. & SOHN, B. C. 2013. Real earnings management and cost of capital. *Journal of Accounting and Public Policy,* 32**,** 518-543.

KIM, J. W. & SHI, Y. 2011. Voluntary disclosure and the cost of equity capital: Evidence from management earnings forecasts. *Journal of Accounting and Public Policy,* 30**,** 348-366.

KLEIN, A. 1996. Can investors use the prospectus to price initial public offerings? *The journal of financial statement analysis,* 2**,** 23-39.

KLINE, R. B. 2011. *Principles and practice of structural equation modeling,* New York, Guilford Press.

KONCHITCHKI, Y., LUO, Y., MA, M. L. Z. & WU, F. 2016. Accounting-based downside risk, cost of capital, and the macroeconomy. *Review of Accounting Studies,* 21**,** 1-36.

KOTHARI, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics,* 31**,** 105-231.

KOTHARI, S. P., LI, X. & SHORT, J. E. 2009. The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review,* 84**,** 1639-1670.

KOTHARI, S. P., SO, E. & VERDI, R. 2016. Analysts' Forecasts and Asset Pricing: A Survey. *In:* LO, A. W. & MERTON, R. C. (eds.) *Annual Review of Financial Economics, Vol 8.*

KRISTANDL, G. & BONTIS, N. 2007. The impact of voluntary disclosure on cost of equity capital estimates in a temporal setting. *Journal of Intellectual Capital,* 8**,** 577-594.

KROSS, W. & SCHROEDER, D. A. 1989. Firm Prominence and the Differential Information Content of Quarterly Earnings Announcements. *Journal of Business Finance & Accounting,* 16**,** 55-74.

KUMAR, P., SORESCU, S. M., BOEHME, R. D. & DANIELSEN, B. R. 2008. Estimation Risk, Information, and the Conditional CAPM: Theory and Evidence. *Review of Financial Studies,* 21**,** 1037-1075.

KYLE, A. S. 1985. Continuous Auctions and Insider Trading. *Econometrica,* 53**,** 1315-1335.

LA PORTA, R. 1996. Expectations and the Cross-Section of Stock Returns. *The Journal of Finance,* 51**,** 1715-1742.

LA PORTA, R., LOPEZ-DE-SILANES, F. & SHLEIFER, A. 2006. What Works in Securities Laws? *The Journal of Finance,* 61**,** 1-32.

LA ROSA, F. & LIBERATORE, G. 2014. Biopharmaceutical and chemical firms' R&D disclosure, and cost of equity: The impact of the regulatory regime. *European Management Journal,* 32**,** 806-820.

LAMBERT, R., LEUZ, C. & VERRECCHIA, R. E. 2007. Accounting Information, Disclosure, and the Cost of Capital. *Journal of Accounting Research,* 45**,** 385-420.

LAMBERT, R. A., LEUZ, C. & VERRECCHIA, R. E. 2012. Information Asymmetry, Information Precision, and the Cost of Capital. *Review of Finance,* 16**,** 1-29.

LAMBERT, R. A. & VERRECCHIA, R. E. 2015. Information, Illiquidity, and Cost of Capital. *Contemporary Accounting Research,* 32**,** 438-454.

LANDSMAN, W. R., MILLER, B. L., PEASNELL, K. & YEH, S. 2011. Do investors understand really dirty surplus? *The Accounting Review,* 86**,** 237-258.

LANG, M. & LUNDHOLM, R. 1993. Cross-sectional determinants of analyst ratings of corporate disclosures. *Journal of Accounting Research***,** 246-271.

LANG, M. H. & LUNDHOLM, R. J. 1996. Corporate Disclosure Policy and Analyst Behavior. *The Accounting Review,* 71**,** 467-492.

LARSON, C. R. & RESUTEK, R. J. 2015. Sources of investor uncertainty and expected stock returns. Available at SSRN: https://ssrn.com/abstract=2318062.

LATIFF, A. R. & TAIB, F. 2011. The Cost of Equity Effect of Accruals Quality and Ownership Structure. *Academy of Taiwan Business Management Review,* 7**,** 1-10.

LEE, C., SO, E. C. & WANG, C. C. 2015. Evaluating firm-level expected-return proxies. *Harvard Business School Accounting & Management Unit Working Paper***,** 15-57.

LEE, C. M., SO, E. C. & WANG, C. C. 2010. Evaluating Implied Cost of Capital Estimates. *SSRN eLibrary*.

LEE, E., STRONG, N. & ZHU, Z. 2014. Did Regulation Fair Disclosure, SOX, and Other Analyst Regulations Reduce Security Mispricing? *Journal of Accounting Research,* 52**,** 733-774.

LEE, P., STOKES, D., TAYLOR, S. & WALTER, T. 2003. The association between audit quality, accounting disclosures and firm-specific risk: Evidence from initial public offerings. *Journal of Accounting and Public Policy,* 22**,** 377-400.

LEVI, S. & ZHANG, X. J. 2015. Do Temporary Increases in Information Asymmetry Affect the Cost of Equity? *Management Science,* 61**,** 354-371.

LEWELLEN, J. & SHANKEN, J. 2002. Learning, Asset-Pricing Tests, and Market Efficiency. *The Journal of Finance,* 57**,** 1113-1145.

LI, K. & MOHANRAM, P. 2014. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies,* 19**,** 1152-1185.

LI, K. K. 2011. How well do investors understand loss persistence? *Review of Accounting Studies,* 16**,** 630-667.

LI, X. 2015. Accounting Conservatism and the Cost of Capital: An International Analysis. *Journal of Business Finance & Accounting,* 42**,** 555-582.

LINTNER, J. 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics,* 47**,** 13-37.

LITTLE, R. J. A. & RUBIN, D. B. 2002. *Statistical analysis with missing data,* Hoboken, N.J., Hoboken, N.J. : Wiley.

LITTLE, T. D., LINDENBERGER, U. & NESSELROADE, J. R. 1999. On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods,* 4**,** 192-211.

LIU, J., NISSIM, D. & THOMAS, J. 2002. Equity Valuation Using Multiples. *Journal of Accounting Research,* 40**,** 135-172.

LIU, L. X., SHERMAN, A. E. & ZHANG, Y. 2014. The Long-Run Role of the Media: Evidence from Initial Public Offerings. *Management Science,* 60**,** 1945-1964.

LIU, M. & WYSOCKI, P. 2016. Cross-Sectional Determinants of Information Quality Proxies and Cost of Capital Measures. *Quarterly Journal of Finance,* 0**,** 1650016.

LOPES, A. B. & DE ALENCAR, R. C. 2010. Disclosure and cost of equity capital in emerging markets: The Brazilian case. *International Journal of Accounting,* 45**,** 443-464.

LOWRY, M., OFFICER, M. S. & SCHWERT, G. W. 2010. The Variability of IPO Initial Returns. *The Journal of Finance,* 65**,** 425-465.

LOWRY, M. & SCHWERT, G. W. 2004. Is the IPO pricing process efficient? *Journal of Financial Economics,* 71**,** 3-26.

LUONG, L. H., NGUYEN, L. H. & YIN, X. 2011. Institutional Investors, Private Information, and the Cost of Capital. Available at SSRN: https://ssrn.com/abstract=1972657.

MADHAVAN, A., RICHARDSON, M. & ROOMANS, M. 1997. Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks. *Review of Financial Studies,* 10**,** 1035-1064.

MANGENA, M., LI, J. & TAURINGANA, V. 2016. Disentangling the Effects of Corporate Disclosure on the Cost of Equity Capital: A Study of the Role of Intellectual Capital Disclosure. *Journal of Accounting Auditing and Finance,* 31**,** 3-27.

MASHRUWALA, C. A. & MASHRUWALA, S. D. 2011. The pricing of accruals quality: January versus the rest of the year. *The Accounting Review,* 86**,** 1349-1381.

MCINNIS, J. 2010. Earnings Smoothness, Average Returns, and Implied Cost of Equity Capital. *The Accounting Review,* 85**,** 315-341.

MCNICHOLS, M. & O'BRIEN, P. C. 1997. Self-Selection and Analyst Coverage. *Journal of Accounting Research,* 35**,** 167-199.

MCNICHOLS, M. F. 2002. The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors: Discussion. *The Accounting Review,* 77**,** 61-69.

MERTON, R. C. 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance,* 42**,** 483-510.

MILLER, E. M. 1977. Risk, uncertainty, and divergence of opinion. *The Journal of Finance,* 32**,** 1151-1168.

MOHANRAM, P. & GODE, D. 2013. Removing predictable analyst forecast errors to improve implied cost of equity estimates. *Review of Accounting Studies,* 18**,** 443-478.

MOHANRAM, P. & RAJGOPAL, S. 2009. Is PIN priced risk? *Journal of Accounting and Economics,* 47**,** 226-243.

MØLLER, A. P. & JENNIONS, M. D. 2001. Testing and adjusting for publication bias. *Trends in Ecology & Evolution,* 16**,** 580-586.

MOSSIN, J. 1966. Equilibrium in a Capital Asset Market. *Econometrica,* 34**,** 768-783.

MUINO, F. & TROMBETTA, M. 2009. Does graph disclosure bias reduce the cost of equity capital? *Accounting and Business Research,* 39**,** 83-102.

NG, A. C. & REZAEE, Z. 2012. Sustainability disclosures and cost of capital. Available at SSRN: https://ssrn.com/abstract=2038654.

NG, A. C. & REZAEE, Z. 2015. Business sustainability performance and cost of equity capital. *Journal of Corporate Finance,* 34**,** 128-149.

O'HARA, M. 1995. *Market microstructure theory,* Blackwell Cambridge, MA.

OGNEVA, M. 2012. Accrual Quality, Realized Returns, and Expected Returns: The Importance of Controlling for Cash Flow Shocks. *The Accounting Review,* 87**,** 1415-1444.

OGNEVA, M., SUBRAMANYAM, K. R. & RAGHUNANDAN, K. 2007. Internal control weakness and cost of equity: Evidence from SOX section 404 disclosures. *The Accounting Review,* 82**,** 1255-1297.

OHLSON, J. A. 1995. Earnings, Book Values, and Dividends in Equity Valuation*. *Contemporary Accounting Research,* 11**,** 661-687.

OHLSON, J. A. & JUETTNER-NAUROTH, B. E. 2005. Expected EPS and EPS growth as determinants of value. *Review of accounting studies,* 10**,** 349-365.

ORENS, R., AERTS, W. & CORMIER, D. 2010. Web-Based Non-Financial Disclosure and Cost of Finance. *Journal of Business Finance & Accounting,* 37**,** 1057-1093.

ORENS, R., AERTS, W. & LYBAERT, N. 2009. Intellectual capital disclosure, cost of finance and firm value. *Management Decision,* 47**,** 1536-1554.

ORENS, R., AERTS, W. & LYBAERT, N. 2013. Customer value disclosure and cost of equity capital. *Review of Accounting and Finance,* 12**,** 130-147.

OTHMAN, R. 2012. Impact of financial reporting quality on the implied cost of equity capital: Evidence from the Malaysian listed firms. *Asian Journal of Business and Accounting,* 3**,** 1-25.

PÁSTOR, Ľ. & STAMBAUGH, ROBERT F. 2003. Liquidity Risk and Expected Stock Returns. *Journal of Political Economy,* 111**,** 642-685.

PAUGAM, L. & RAMOND, O. 2015. Effect of Impairment-Testing Disclosures on the Cost of Equity Capital. *Journal of Business Finance & Accounting,* 42**,** 583-618.

PEARLMAN, K., SCHMIDT, F. L. & HUNTER, J. E. 1980. Validity generalization results for tests used to predict job proficiency and training success in clerical occupations. *Journal of Applied Psychology,* 65**,** 373.

PHILBRICK, D. R. & RICKS, W. E. 1991. Using value line and IBES analyst forecasts in accounting research. *Journal of Accounting Research***,** 397-417.

PLUMLEE, M., BROWN, D., HAYES, R. M. & MARSHALL, R. S. 2015. Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy,* 34**,** 336-361.

POSHAKWALE, S. & COURTIS, J. K. 2005. Disclosure level and cost of equity capital: evidence from the banking industry. *Managerial and Decision Economics,* 26**,** 431-444.

RAO, P. 1973. Some Notes on the Errors-in-Variables Model. *The American Statistician,* 27**,** 217-218.

RESUTEK, R. J. 2011. Discussion of "How well do investors understand loss persistence?". *Review of Accounting Studies,* 16**,** 668-678.

REVERTE, C. 2012. The Impact of Better Corporate Social Responsibility Disclosure on the Cost of Equity Capital. *Corporate Social Responsibility and Environmental Management,* 19**,** 253-272.

RICHARDSON, A. J. & WELKER, M. 2001. Social disclosure, financial disclosure and the cost of equity capital. *Accounting, Organizations and Society,* 26**,** 597-616.

RICHARDSON, S. A., SLOAN, R. G., SOLIMAN, M. T. & TUNA, I. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics,* 39**,** 437-485.

RITTER, J. R. 1991. The Long-Run Performance of initial Public Offerings. *The Journal of Finance,* 46**,** 3-27.

RITTER, J. R. & WELCH, I. 2002. A review of IPO activity, pricing, and allocations. *The Journal of Finance,* 57**,** 1795-1828.

ROSENTHAL, R. 1991. *Meta-analytic procedures for social research,* Sage.

ROSS, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory,* 13**,** 341-360.

RUBIN, D. B. 1976. Inference and missing data. *Biometrika,* 63**,** 581-592.

RUUD, J. S. 1993. Underwriter price support and the IPO underpricing puzzle. *Journal of Financial Economics,* 34**,** 135-151.

SAFDAR, R. & YAN, C. 2016. Information risk, stock returns, and the cost of capital in China. *China Finance Review International,* 6**,** 77-95.

SAINI, J. S. & HERRMANN, D. 2012. Cost of equity capital, information asymmetry, and segment disclosure. Available at SSRN: https://ssrn.com/abstract=1758934.

SAVALEI, V. & BENTLER, P. M. 2010. Structural Equation Modeling. *The Corsini Encyclopedia of Psychology.* John Wiley & Sons, Inc.

SCHAFER, J. L. & GRAHAM, J. W. 2002. Missing data: Our view of the state of the art. *Psychological Methods,* 7**,** 147-177.

SHARPE, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance,* 19**,** 425-442.

SHENG, X. & THEVENOT, M. 2012. A new measure of earnings forecast uncertainty. *Journal of Accounting and Economics,* 53**,** 21-33.

SHENG, X. & THEVENOT, M. 2015. Quantifying differential interpretation of public information using financial analysts' earnings forecasts. *International Journal of Forecasting,* 31**,** 515-530.

SHEVLIN, T. 2013. Some personal observations on the debate on the link between financial reporting quality and the cost of equity capital. *Australian Journal of Management,* 38**,** 447-473.

SHLEIFER, A. 1986. Do Demand Curves for Stocks Slope Down? *The Journal of Finance,* 41**,** 579-590.

SOUISSI, M. & KHLIF, H. 2012. Meta-analytic review of disclosure level and cost of equity capital. *International Journal of Accounting & Information Management,* 20**,** 49-62.

STOLL, H. R. 1978. The Pricing of Security Dealer Services: An Empirical Study of Nasdaq Stocks. *The Journal of Finance,* 33**,** 1153-1172.

STROBL, G. 2013. Earnings Manipulation and the Cost of Capital. *Journal of Accounting Research,* 51**,** 449-473.

STRONG, N. 1993. The Relation between Returns and Earnings: Evidence for the UK. *Accounting and Business Research,* 24**,** 69-77.

STRONG, N. & WALKER, M. 1993. The Explanatory Power of Earnings for Stock Returns. *The Accounting Review,* 68**,** 385-399.

TASKER, S. C. 1998. Bridging the information gap: Quarterly conference calls as a medium for voluntary disclosure. *Review of Accounting Studies,* 3**,** 137-167.

TOHANG, V. & HUTAGAOL-MARTOWIDJOJO, Y. R. I. 2015. Corporate Voluntary Disclosure and Cost of Equity in an Emerging Country: Evidence from Indonesia. *Advanced Science Letters,* 21**,** 898-902.

UNGER, L. S. 2000. *Speech by SEC Commissioner: Fallout from Regulation FD Has the SEC Finally Cut the Tightrope?* [Online]. Available: https://www.sec.gov/news/speech/spch421.htm [Accessed 01/04/2016].

VENTER, J. & DE JONGH, D. 2006. Extending the EKOP model to estimate the probability of informed trading. *Studies in Economics and Econometrics,* 30**,** 25-39.

VUOLTEENAHO, T. 2002. What Drives Firm-Level Stock Returns? *The Journal of Finance,* 57**,** 233-264.

WAYMIRE, G. 1985. Earnings volatility and voluntary management forecast disclosure. *Journal of Accounting Research***,** 268-295.

WELKER, M. 1995. Disclosure Policy, Information Asymmetry, and Liquidity in Equity Markets. *Contemporary Accounting Research,* 11**,** 801-827.

WILLIAM LIN, H.-W. & KE, W.-C. 2011. A computing bias in estimating the probability of informed trading. *Journal of Financial Markets,* 14**,** 625-640.

WU, S.-W., LIN, F. & WU, C.-M. 2014. Corporate Social Responsibility and Cost of Capital: An Empirical Study of the Taiwan Stock Market. *Emerging Markets Finance and Trade,* 50**,** 107-120.

XIAO-FENG, H., WEI-LING, L. & MING-YI, L. Empirical Study of the Relationship between Voluntary Disclosure and the Cost of Equity Capital. Management Science and Engineering, 5-7 Oct. 2006 2006. 1446-1451.

XU, R. Z. 2009. Reexamination of the relationship between disclosure and cost of capital. *Global Journal of Business Research,* 3**,** 15-26.

YAN, Y. & ZHANG, S. 2012. An improved estimation method and empirical properties of the probability of informed trading. *Journal of Banking & Finance,* 36**,** 454-467.

YAN, Y. & ZHANG, S. 2014. Quality of PIN estimates and the PIN-return relationship. *Journal of Banking & Finance,* 43**,** 137-149.

ZHANG, X. F. 2006. Information Uncertainty and Stock Returns. *The Journal of Finance,* 61**,** 105-137.

ZHAO, Y., DAVIS, M. & BERRY, K. T. 2009. Disclosure channel and cost of capital: evidence from open vs closed conference calls. *Review of Accounting and Finance,* 8**,** 253-278.